

Global Scheduling of Multi-Mode Real-Time Applications upon Multiprocessor Platforms

Vincent Nélis

Patrick Meumeu Yomsi

Björn Andersson

Joël Goossens

Abstract

Multi-mode real-time systems are those which support applications with different modes of operation, where each mode is characterized by a specific set of tasks. At run-time, such systems can, at any time, be requested to switch from its current operating mode to another mode (called “new mode”) by replacing the current set of tasks with that of the new-mode. Thereby, ensuring that all the timing requirements are met not only requires that a schedulability test is performed on the tasks of each mode but also that (i) a protocol for transitioning from one mode to another is specified and (ii) a schedulability test for each transition is performed. We propose two distinct protocols that manage the mode transitions upon uniform and identical multiprocessor platforms at run-time, each specific to distinct task requirements. For each protocol, we formally establish schedulability analyses that indicate beforehand whether all the timing requirements will be met during any mode transition of the system. This is performed assuming both Fixed-Task-Priority and Fixed-Job-Priority schedulers.

1 Introduction

Hard real-time systems require both functionally correct executions and *results that are produced on time*. Control of the traffic (ground or air), control of engines, control of chemical and nuclear power plants are just some examples of such systems. Currently, numerous techniques exist that enable engineers to design real-time systems while guaranteeing that all the temporal requirements are met. These techniques generally model each functionality of the application by a *recurrent* task, characterized by a computing requirement, a temporal deadline and an activation rate. Commonly, real-time applications are simply modeled by a single and finite set of such tasks. However, practical applications often exhibit multiple behaviors issued from several operating modes (e.g., an initialization mode, an emergency mode, a fault recovery mode, etc.), where each mode is characterized by its own set of functionalities, i.e., its set of tasks. During the execution of such *multi-mode* real-time applications, switching from the current mode (called the *old-mode*) to any other mode (called the *new-mode*) requires to substitute the currently executing task set with the set of tasks of the new-mode. This substitution introduces a *transient phase*, where tasks of both the old- and new-mode may be scheduled *simultaneously*, thereby leading to a possible overload that can compromise the system schedulability—indeed it can be the case that both the old- and new-mode have been asserted schedulable by the schedulability analysis but the transition between them fails at run-time.

The scheduling problem during a transition between two modes has multiple aspects, depending on the behavior and requirements of the old- and new-mode tasks when a mode change is initiated. Upon a mode change request:

- an *old-mode* task may be allowed to be immediately aborted or, on the contrary, can be required to complete the execution of its current active job (so that it preserves data consistency for instance). Using scheduling algorithms such as the one considered in this study, we will prove in Section 5 that aborting tasks upon a mode change request does not jeopardize the schedulability of the mode transitions. Hence, we assume in this paper the most problematic scenario in which *every old-mode task must complete its current active job* (if any) when a mode change is requested.
- a *new-mode* task either requires to be activated as soon as possible when a mode change is requested or requires to be activated only when all the active jobs issued from the old-mode have totally completed their execution.

Finally, there may be some tasks (called mode-independent tasks in the literature) that belong to more than one mode and such that their activation pattern must not be jeopardized during the transition between those modes¹. However this paper will consider only systems that do *not* include such tasks.

¹In practice, mode-independent tasks typically allow to model daemon functionalities.

Transition scheduling protocols for tasks without mode-independent tasks are often classified with respect to the way they schedule the old- and new-mode tasks during the transitions. In the literature (see for instance [34] which considers uniprocessor systems), the following definitions are used.

Definition 1 (Synchronous/asynchronous protocol [34]) A transition protocol is said to be *synchronous* if it schedules new-mode tasks only when all the old-mode tasks have completed. Otherwise, it is said to be *asynchronous*.

Definition 2 (Protocol with/without periodicity [34]) A transition protocol is said to be “*with periodicity*” if and only if it is able to deal with mode-independent tasks. Otherwise, it is said to be “*without periodicity*”.

1.1 Related work

Numerous transition protocols have been proposed for *uniprocessor* platforms (a survey about this concern is presented in [34]). In such environments, existing researches [34, 25, 33] have shown that even if two modes of the application have been proven feasible, the transition between the two modes can cause violation of timing constraints, hence needing explicit analyses. Such analyses have been proposed in [35], considering the popular Rate Monotonic Algorithm. Unfortunately three years later, this analysis was shown optimistic [39] in the sense that some unfeasible task sets could be asserted schedulable. In the same paper [39], the authors improved the previous analysis and proposed a new one which considers the popular Deadline Monotonic Algorithm. An analysis of sporadic tasks scheduled on EDF is known as well [2]. In [37], the authors proposed an analysis which considers Fixed-Task-Priority scheduling (FTP), Earliest-Deadline-First [26] scheduling and arbitrary task activation pattern. Furthermore, for applications that were initially proven not schedulable during the transition phases, they derived the required offsets for delaying the initialization of transition between two modes in order to make the application schedulable.

Among the uniprocessor *synchronous* protocols, the authors of [3, 38, 34] proposed the following protocols.

- ▷ The *Minimum Single Offset Protocol* (MSO) [34] where the last activation of each old-mode task completes and then, the new-mode tasks are released.
- ▷ The *Idle Time Protocol* (IT) [38] where the periodic activations of the old-mode tasks are suspended at the first idle time-instant occurring during the transition and then, the new-mode tasks are released.
- ▷ The *Maximum-Period Offset Protocol* (MPO) [3] where the delays of first activation of each new-mode task is equal to the period of the less frequent task in both modes,

Among the uniprocessor *asynchronous* protocols, the authors of [39, 32, 2] proposed the following protocols.

- ▷ A protocol *without periodicity* [32] where tasks are assigned priorities according to the Deadline Monotonic Scheduling algorithm and are scheduled with time offsets during the mode change only.
- ▷ A protocol *with periodicity* has been introduced by Sha et al. in [36], assuming Fixed-Task-Priority scheduling. Then, the authors of [2] extended this protocol to the Earliest Deadline First [26] scheduling algorithm.
- ▷ The authors of [39] introduced a particular protocol which allows tasks to modify their parameters (period, execution time, etc.) during the mode changes. As in [32], this study assumes that the tasks are scheduled according to the Deadline Monotonic scheduling algorithm.

1.2 Contribution and paper organization

In this paper we propose two protocols *without periodicity* (SM-MSO which is *synchronous* and AM-MSO which is *asynchronous*) for managing mode transitions during the execution of multi-mode real-time applications on *multiprocessor* platforms. Both protocols can be considered as a generalization to multiprocessors of the MSO protocol proposed in [34]. We assume that every operating mode of the application is scheduled by a *global*, *work-conserving*, *preemptive* and *Fixed-Job-Priority* (FJP) scheduling algorithm (formal definitions are given in Section 2.4). Some of the results presented here have already been published (see [27, 29, 31, 30]). It is worth noticing that the problem of scheduling multi-mode applications upon multiprocessor platforms is much more complex than upon uniprocessor platforms, especially due to the presence of scheduling anomalies (see Chapter 5 of [1] for a definition) and it is now well known that real-time multiprocessor scheduling problems are typically not solved by applying straightforward extensions of techniques used for solving similar uniprocessor problems.

The paper is organized as follows. Section 2 defines the computational model used throughout the paper. Sections 3 and 4 describe the synchronous and asynchronous protocols SM-MSO and AM-MSO, respectively. Section 5 introduces some definitions and basic results necessary for the establishment of our schedulability

analyses. These four first Sections 2–5 are a common base of the paper, in the sense that these 20 pages describe both the models of computation and protocols independently of the platform and scheduler characteristics. Then, the four next Sections 6–9 are each specific to a platform and scheduler model. More precisely, they provide a schedulability analysis for both SM-MSO and AM-MSO, assuming in turn identical platforms and Fixed-Job-Priority schedulers (in Section 6), identical platforms and Fixed-Task-Priority schedulers (in Section 7), uniform platforms and Fixed-Job-Priority schedulers (in Section 8) and uniform platforms and Fixed-Task-Priority schedulers (in Section 9)². Finally, Section 10 gives our conclusions and future work, together with some remaining open problems.

2 Models of computation and specifications

2.1 Application specifications

We define a multi-mode real-time application τ as a set of x operating modes denoted by M^1, M^2, \dots, M^x where each mode M^k has to execute its associated task set $\tau^k \stackrel{\text{def}}{=} \{\tau_1^k, \tau_2^k, \dots, \tau_{n_k}^k\}$ composed of n_k tasks by following the scheduler \mathcal{S}^k . At run-time, the application is either running in one and only one mode, i.e., it is executing only the set of tasks associated to that mode, or it is switching from one mode to another one. Since we do not consider mode-independent tasks in this study, it holds that $\tau^k \cap \tau^j = \emptyset, \forall k \neq j$.

Each task τ_i^k is modeled by a *sporadic* and *constrained-deadline* task characterized by three parameters $\langle C_i^k, D_i^k, T_i^k \rangle$ —a worst-case execution time C_i^k , a minimum inter-arrival time T_i^k and a relative deadline $D_i^k \leq T_i^k$ —with the interpretation that, during the execution in mode M^k , task τ_i^k generates successive jobs $\tau_{i,j}^k$ (with $j = 1, \dots, \infty$) released at times $a_{i,j}^k$ such that $a_{i,j}^k \geq a_{i,j-1}^k + T_i^k$ (with $a_{i,1}^k \geq 0$), each such job has an execution requirement of at most C_i^k , and must be completed at (or before) its absolute deadline noted $d_{i,j}^k \stackrel{\text{def}}{=} a_{i,j}^k + D_i^k$. In the particular case where $a_{i,j}^k = a_{i,j-1}^k + T_i^k, \forall j > 1$, the task τ_i^k is said to be *periodic* instead of *sporadic*. In the same vein, if $D_i^k = T_i^k$ then the task τ_i^k is said to be *implicit-deadline* instead of *constrained-deadline*.

Definition 3 (Active job) We say that a job $\tau_{i,j}^k$ is active at time t if it has been already released (i.e., $t \leq a_{i,j}^k$) and it is not completed yet.

Since we assume $D_i^k \leq T_i^k$, there cannot be two jobs of a same task τ_i^k active at a same time in any feasible schedule. All the tasks are assumed to be independent, i.e., there is no communication, no precedence constraint and no shared resource (except the processors) between them. In [31], we introduced the following concept of enabled/disabled tasks.

Definition 4 (Enabled/disabled tasks [31]) At run-time, any task τ_k^i of the application can generate jobs if and only if τ_k^i is enabled. Symmetrically, a disabled task cannot generate jobs.

As such, disabling a task τ_k^i prevents future job releases from τ_k^i . When all the tasks of any mode τ^i are enabled and all the tasks of all the other modes are disabled, the application is said to be running in mode M^i (since only the tasks of mode τ^i can release jobs). We denote by $\text{enabled}(\tau^i, t)$ and $\text{disabled}(\tau^i, t)$ the subsets of enabled and disabled tasks of τ^i at time t , respectively.

2.2 Platform specifications

Many recent embedded systems are built upon multiprocessor platforms in order to fulfill the high computational requirements of applications. As pointed out in [8, 9], another advantage of such a choice is the fact that multiprocessor systems are more energy efficient than equally powerful uniprocessor platforms. Indeed, raising the frequency of a single CPU results in a *multiplicative* increase of the consumption while adding CPUs results in an *additive* increase. Two distinct multiprocessor architectures are commonly used in the industrial world and thus, are considered in this paper: identical and uniform platforms.

Identical platform. In such multiprocessor platforms, all the CPUs have the same computational capabilities, with the interpretation that in any interval of time two CPUs execute the same amount of work (assuming that none of them is idling). In the remainder of this paper, any platform composed of m identical CPUs will be modeled by $\pi \stackrel{\text{def}}{=} \{\pi_1, \pi_2, \dots, \pi_m\}$ where π_i denotes the i^{th} CPU of the platform.

Uniform platform. In such multiprocessor platforms, the CPUs are allowed to have different computational capabilities. That is, a parameter s_i is associated to every CPU π_i with the interpretation that in any time interval of length t , CPU π_i executes $s_i \cdot t$ units of execution (if it is not idling). This parameter can be seen

²Even though Fixed-Job-Priority schedulers encompass the family of Fixed-Task-Priority schedulers, the particular case of Fixed-Task-Priority schedulers is treated separately so that the schedulability analyses are more specific and therefore more accurate.

as the *execution speed* of the CPU. In the remainder of this paper, any platform composed of m uniform CPUs is modeled by $\pi \stackrel{\text{def}}{=} \{s_1, s_2, \dots, s_m\}$, where s_i is the execution speed of CPU π_i . Without loss of generality, we assume that $s_i \geq s_{i-1} \forall i = 2, 3, \dots, m$, meaning that CPU π_m is the fastest CPU while π_1 is the slowest one. For all $k \in [1, m]$, we denote by $s(k)$ the cumulated speed of the $(m - k + 1)$ fastest CPUs, i.e.,

$$s(k) \stackrel{\text{def}}{=} \sum_{i=k}^m s_i \quad (1)$$

Notice that identical platforms are a particular case of uniform platforms where $s_i = s_j \forall i, j \in [1, m]$. In this particular case we assume without any loss of generality that $\forall i: s_i = 1$.

2.3 Mode transition specifications

While the application is running in any mode M^i , a mode change can be initiated by any task of τ^i or by the system itself, whenever it detects a change in the environment or in its internal state for instance. This is performed by invoking a $\text{MCR}(j)$ (i.e., a Mode Change Request), where M^j is the destination mode. We denote by $t_{\text{MCR}(j)}$ the invoking time of the *last* $\text{MCR}(j)$. From the time at which a mode change is requested to the time at which the transition phase ends, M^i and M^j are referred to as the old- and new-mode, respectively.

At run-time, mode transitions are managed as follows. Suppose that the application is running in mode M^i and the system (or any task of τ^i) comes to request a mode change to mode M^j , with $j \neq i$. At time $t_{\text{MCR}(j)}$, the system entrusts the scheduling decisions to a transition protocol which *immediately* disables all the old-mode tasks, thus preventing them from releasing new jobs. At this time, the active jobs issued from these disabled tasks, henceforth called the *rem-jobs* (for “remaining jobs”), may have two distinct behaviors: either *they can be aborted* upon the $\text{MCR}(j)$ or *they can complete their execution*. From the schedulability point of view, we will show that aborting some (or all) rem-jobs upon a mode change request does not jeopardize the system schedulability during the transition phase. Consequently, we assume the worst-case scenario for every mode transition, i.e., the scenario in which *every old-mode task has to complete its last released job (if any) during every mode transition*³. The fact that the rem-jobs have to complete their execution upon the $\text{MCR}(j)$ brings the following problem. Even if both task sets τ^i and τ^j (from the old- and new-mode, respectively) have been asserted to be schedulable upon the m CPUs at system design-time, the presence of the rem-jobs may cause an *overload* during the transition phase (at run-time) if all the new-mode tasks of τ^j are enabled *immediately* upon the mode change request. Indeed, the schedulability analysis performed beforehand on τ^j did not take into account the additional work generated by the rem-jobs. To solve this problem, transition protocols usually delay the enablement of each new-mode task until it is safe to do so. However, these delays are also subject to hard constraints. More precisely, we denote by $\mathcal{D}_k^j(M^i)$ the *relative transition deadline* of task τ_k^j during every transition from mode M^i to mode M^j , with the following interpretation: the transition protocol must ensure that τ_k^j is enabled not later than time $t_{\text{MCR}(j)} + \mathcal{D}_k^j(M^i)$. Finally, when all the rem-jobs are completed and all the new-mode tasks of τ^j are enabled, the system entrusts the scheduling decisions to the scheduler S^j of the new-mode M^j and the transition phase ends.

In short, the goal of any transition protocol is to fulfill the following requirements during every mode change:

1. Complete each rem-job $\tau_{a,b}^i$ by its absolute deadline $d_{a,b}^i$.
2. Enable each new-mode task τ_k^j by its absolute transition deadline $t_{\text{MCR}(j)} + \mathcal{D}_k^j(M^i)$.
3. Complete each new-mode job⁴ $\tau_{a,b}^j$ by its absolute deadline $d_{a,b}^j$.

Definition 5 (Valid protocol [31]) A transition protocol \mathcal{A} is said to be valid for a given application τ and platform π if and only if \mathcal{A} meets all the job and transition deadlines during every transition between every pair of operating modes of τ .

This notion of “valid protocol” is directly related to that of a “validity test” defined as follows.

Definition 6 (Validity test [31]) For a given transition protocol \mathcal{A} , a validity test is a condition based on the tasks and platform characteristics that indicates a priori whether \mathcal{A} is valid for a given application τ and platform π .

³Aborting a job consists in suddenly stopping its execution and removing it from the system memory. But in the real world, suddenly killing a process may cause system failures and the rem-jobs often have to complete their execution.

⁴This requirement is automatically fulfilled for synchronous protocols since no new-mode jobs are scheduled during the mode transitions.

2.4 Scheduler specifications

We consider the *global* preemptive scheduling problem of sporadic constrained-deadline tasks upon multiprocessor platforms. “Global” schedulers, in contrast to partitioned ones, allow different tasks and *different* jobs of the same task to be executed upon *different* CPUs. When preemptive, global schedulers allow any job to be interrupted at any time prior to completion on any CPU and resumed (possibly later) on any other CPU. We consider that every mode M^k uses its own scheduler denoted by S^k which can be either *Fixed-Task-Priority (FTP)* or *Fixed-Job-Priority (FJP)* according to the following interpretations.

- FTP schedulers assign a priority to each task at system design-time (i.e., before the execution of the application) and then at run-time, every released job uses the priority of its task and the priority of a job is kept constant until it completes.
- FJP schedulers assign a priority to each job at run-time (i.e., as soon as it arrives in the system) and every job keeps its priority constant until it completes. As such, different jobs issued from the same task may have different priorities⁵.

Without loss of generality we assume that, at any time, two active jobs cannot have the same priority. Furthermore, we consider *work-conserving* schedulers according to the following definition.

Definition 7 (Work-conserving global scheduler) A CPU cannot be idle if there is a job awaiting execution. Usually, priority-based schedulers assign at each instant in time the m highest priority active jobs (if any) to the m CPUs.

The above definition of work-conserving schedulers encompasses a large family of schedulers, but suffers from an important lack of *determinism*. Indeed for a given set of jobs, multiple (and different) schedules can sometimes be derived from the same work-conserving scheduler (and thus from the same job priority assignment). The following example illustrates this drawback.

Example 1 Let us consider the set J of 5 jobs with respective processing time 4, 8, 4, 4 and 6. Suppose that J is scheduled on a 2-processors identical platform π by an FTP, global, preemptive and work-conserving scheduler such that $J_1 > J_2 > J_3 > J_4 > J_5$. According to Definition 7, Figures 1 and 2 depict two possible different schedules corresponding to this priority assignment.

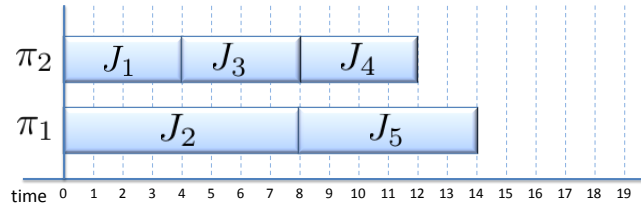


Figure 1: A possible schedule of J_1, J_2, J_3, J_4 and J_5 .

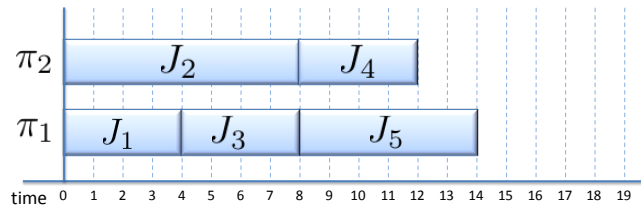


Figure 2: Another possible schedule of J_1, J_2, J_3, J_4 and J_5 .

In order to get around this lack of determinism, we introduce two refinements of Definition 7 that we name *weakly* and *strongly* work-conserving schedulers, respectively. Weakly work-conserving schedulers concern only identical platforms whereas strongly work-conserving schedulers concern only in uniform (and non-identical) platforms. The rationale for introducing these two refinements is to have *one and only one possible schedule* for any given set of synchronous⁶ jobs, multiprocessor platform and job priority assignment.

⁵According to these interpretations, FTP schedulers are a particular case of FJP schedulers in which all the jobs issued from a same task receive the same priority determined beforehand.

⁶The term “synchronous” jobs is commonly used in the literature to refer to jobs that are all ready for execution at the same time.

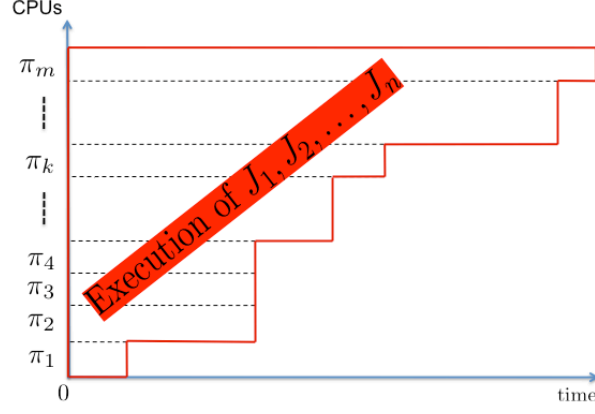


Figure 3: For any fixed set of jobs and uniform platform, the schedule generated by any strongly work-conserving scheduler forms a staircase.

Definition 8 (Weakly work-conserving scheduler) A scheduler S is weakly work-conserving if and only if:

- no CPU idles while there are active jobs awaiting execution, and
- if there are more than one job awaiting execution and more than one CPU available for the execution of those jobs then S assigns the highest priority waiting job to the available CPU with the highest index.

Property 1 (Unique schedule) For any given finite set J of jobs, any weakly work-conserving scheduler S and any identical multiprocessor platform π , there exists one and only one possible schedule of J upon π following S .

In order to illustrate this property, let us consider the set of 5 jobs used in Example 1, a 2-processors identical platform π and any weakly work-conserving scheduler assigning priorities such that $J_1 > J_2 > J_3 > J_4 > J_5$. The unique possible schedule of J upon π is the one depicted in Figure 1. Indeed at time 0, CPUs π_1 and π_2 are idle and the second condition of Definition 8 imposes J_1 to execute on π_2 . From the same rule, J_4 must execute on π_2 at time 8. Notice that the refinement of “weakly” work-conserving scheduler clarifies only the job-to-CPU assignment rule when the highest-priority waiting job has to be dispatched to a CPU.

Definition 9 (Strongly work-conserving scheduler) A scheduler S is strongly work-conserving if and only if:

- no CPU idles while there are active jobs awaiting execution, and
- at every time during the system execution, the job-to-CPU assignment uses the rule: highest priority active job upon highest indexed CPU.

In contrast to the refinement of “weakly” work-conserving schedulers, the “strongly”-refinement clarifies the job-to-CPU assignment rule *at each time-instant* during the system execution. It is *essential* to keep in mind that in our study weakly work-conserving schedulers will be used *only* on *identical* platforms whereas strongly work-conserving schedulers will be used *only* on *uniform* and *non-identical* platforms. For strongly work-conserving schedulers, the concept of migrating jobs to faster CPUs as soon as possible (as specified by the second condition of Definition 9) has been widely used over the years on uniform platforms (see [12, 14, 20, 13, 17, 18]). This refinement is extremely important, especially because it yields the following property.

Property 2 (Staircase property) Let J denote any finite set of synchronous jobs, π any uniform multiprocessor platform and S any strongly work-conserving scheduler. In the schedule of J upon π by S , CPU π_ℓ idles before or at the same time-instant as CPU $\pi_{\ell+1}$ for all $\ell < m$.

Informally speaking, the schedule of J upon π by S forms a *staircase* (see Figure 3).

This property stems from the fact that the CPUs are indexed in such a manner that $s_i \geq s_j \forall i > j$. Thus, it holds from the second condition of Definition 9 that at any instant t , if S idles the i^{th} -slowest CPU then S also idles the j^{th} slowest CPUs for all $j < i$. Also, it results from the same condition that the i^{th} CPU that starts idling is always π_i .

The following definition introduces the fundamental notion of *predictability*, and Lemmas 1 and 2 are essential for the rest of the paper.

Definition 10 (Predictability [24]) Let A denote a scheduler, and let $J = \{J_1, J_2, J_3, \dots\}$ be a potentially infinite set of jobs, where each job $J_i = (a_i, c_i, d_i)$ is characterized by an arrival time a_i , a computing requirement c_i and an absolute deadline d_i . Let r_i and f_i denote the time at which job J_i starts and completes its execution (respectively) when J is scheduled by A . Now, consider any set $J' = \{J'_1, J'_2, J'_3, \dots\}$ of jobs obtained from J as follows. Job J'_i has an arrival time a_i , an execution requirement $c'_i \leq c_i$, and a deadline d_i . Let r'_i and f'_i denote the time at which job J'_i starts and completes its execution (respectively) when J' is scheduled by A . Algorithm A is said to be predictable if and only if for any set of jobs J and for any such J' obtained from J , it is the case that $r'_i \leq r_i$ and $f'_i \leq f_i \forall i$.

Informally speaking, Definition 10 claims that an upper-bound on the starting time and on the completion time of each job can be determined by analyzing the situation under the assumption that each job executes for its WCET. The result from [22, 24, 23] that we will be using can be stated as follows.

Lemma 1 (See [22, 24, 23]) On identical multiprocessor platforms, any FJP, global, preemptive and weakly work-conserving scheduler is predictable.

In the same vein, the result from [17] that we will be using can be stated as follows.

Lemma 2 (See [17]) On uniform multiprocessor platforms, any FJP, global, preemptive and strongly work-conserving scheduler is predictable.

We use the notation \mathcal{P} to refer to a specific job priority assignment. A job priority assignment can be seen as a key component of any scheduler, but the definition of a scheduler is more general since, in addition to a job priority assignment, a scheduler must also provide specifications like “global or partitioned”, “preemptive or non-preemptive”, etc. For any job priority assignment \mathcal{P} , we denote by $J_i >_{\mathcal{P}} J_j$ the fact that job J_i has a higher priority than J_j according to \mathcal{P} , and we assume that every assigned priority is distinct from the others. That is, $\forall \mathcal{P}, i, j$ such that $i \neq j$ we have either $J_i >_{\mathcal{P}} J_j$ xor $J_i <_{\mathcal{P}} J_j$. Similarly, and without any distinction with the interpretation given above, we will sometimes use the notations $J_i >_{S^k} J_j$ and $J_i <_{S^k} J_j$ where S^k is the scheduler of mode M^k , and we will sometimes use the notations $J_i > J_j$ and $J_i < J_j$ when the job priority assignment has no label (for instance, when we will depict some examples of schedules, we will just say “ $J_i > J_j$ ” without giving a name to the job priority assignment). Finally, the problems and solutions presented in this paper are addressed under the following assumptions:

- ▷ **Assumption 1.** The set τ^k of tasks of every mode M^k can be scheduled by S^k on m CPUs without missing any deadline.
- ▷ **Assumption 2.** Job migrations and preemptions are permitted and are carried out at no loss or penalty.
- ▷ **Assumption 3.** Job parallelism is forbidden, i.e., jobs execute on at most one CPU at any instant in time.
- ▷ **Assumption 4.** For every mode M^i it holds that $m \leq n_i$, where n_i is the number of tasks in mode M^i .

Regarding Assumption 1, it allows us to focus only on the schedulability of the application during the *transient phases* corresponding to mode transitions, rather than on the schedulability of the application during the execution in a given mode.

Regarding Assumption 4, it is worth noticing that since job parallelism is forbidden and tasks are assumed to be constrained-deadline, there are at most n_i jobs active at a same time during the execution of any mode M^i . As a result, it holds for each mode M^i that in every schedulable application where $m > n_i$, there are always $m - n_i$ CPUs that constantly idle. We will see later that these $m - n_i$ idling CPUs are the slowest ones and the problem in that case thereby reduces to the same problem upon the subset of the n_i fastest CPUs among these m CPUs.

3 The synchronous protocol SM-MSO

3.1 Description of the protocol

The protocol SM-MSO (which stands for “Synchronous Multiprocessor Minimum Single Offset” protocol) is an extension to multiprocessor platforms of the protocol MSO defined in [34] for uniprocessor platforms. This protocol supports both uniform and identical platforms. The main idea of SM-MSO is the following: upon a $\text{MCR}(j)$, $\forall j$, all the tasks of the old-mode (say M^i) are disabled and the rem-jobs continue to be scheduled by the old-mode scheduler S^i upon the m CPUs. Once *all* the rem-jobs are completed, *all* the new-mode tasks (i.e., the tasks of τ^j) are simultaneously enabled. Algorithm 4 gives the pseudo-code of this protocol and Example 2 illustrates how SM-MSO handles the mode transitions.

Input: M^i : the old mode
Input: M^j : the new-mode
Input: the rem-jobs
1: **while true do**
2: Schedule the rem-jobs according to S^i
3: **if** (any rem-job J_k completes at time t) **then**
4: **if** ($\text{active}(\tau^i, t) = \phi$) **then**
5: enable all the new-mode tasks of τ^j
6: enter the new-mode M^j
7: **end if**
8: **end if**
9: **end while**

Figure 4: SM-MSO protocol

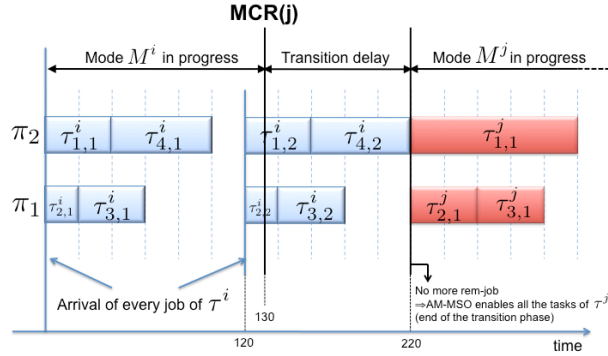


Figure 5: Illustration of a mode transition handled by SM-MSO.

Example 2 Let us consider a platform π composed of only 2 identical CPUs and an application composed of 2 modes M^i and M^j depicted in blue and red, respectively. We assume that these two modes contain only synchronous implicit-deadline periodic tasks. The old-mode M^i contains 4 tasks with characteristics given in Table 1 and uses an FTP scheduler S^i such that $\tau_1^i >_{S^i} \tau_2^i >_{S^i} \tau_3^i >_{S^i} \tau_4^i$.

Tasks	C_k^i	$D_k^i = T_k^i$
τ_1^i	40	120
τ_2^i	20	120
τ_3^i	40	120
τ_4^i	60	120

Table 1: Characteristics of the tasks in M^i .

The new-mode M^j contains 3 tasks $\tau_1^j, \tau_2^j, \tau_3^j$ and uses an FTP scheduler S^j such that $\tau_1^j >_{S^j} \tau_2^j >_{S^j} \tau_3^j$. The characteristics of these tasks are: $C_1^j = 100$ and $C_2^j = C_3^j = 40$. The deadline and period of these new-mode tasks do not have any importance in this example and we intentionally omitted to specify them. Figure 5 illustrates the SM-MSO transition protocol between these two modes.

At time 120, every task of M^i releases its second job and the scheduler S^i starts the execution of $\tau_{1,2}^i$ and $\tau_{2,2}^i$ on CPU π_2 and π_1 , respectively. Then, suppose that the system requests a mode change at time 130. Here starts the transition phase from mode M^i to mode M^j . As specified by the protocol SM-MSO, all the old-mode tasks are immediately disabled and the remaining active jobs $\tau_{1,2}^i, \tau_{2,2}^i, \tau_{3,2}^i$ and $\tau_{4,2}^i$ (named the rem-jobs from this point forward) continue to be scheduled according to the old-mode scheduler S^i . These rem-jobs execute until time 220, time at which they are all completed. At this instant 220, the condition at line 4 of Algorithm 4 is verified. Thus, SM-MSO enables all the new-mode tasks and starts scheduling the incoming new-mode jobs according to the new-mode scheduler S^j . Notice that at any time during every transition phase, our protocol SM-MSO allows the system (or any task) to request any other mode change. At the very end of the current transition phase (at time 220 in this example), SM-MSO enables all the tasks of the mode M^z assuming that $\text{MCR}(z)$ is the last mode change that has been requested.

3.2 Design of a validity test

In order to establish a *validity test* for the protocol SM-MSO, two key results are required:

1. It must be proved for every mode transition that disabling the old-mode tasks upon a MCR does not jeopardize the schedulability of the rem-jobs when they continue to be scheduled by the old-mode scheduler. That is, it must be guaranteed that the absolute deadline $d_{a,b}^i$ of every rem-job $\tau_{a,b}^i$ is met during every mode transition from every mode M^i .
2. It must be proved for every mode transition that the length of the transition phase *can never be larger* than the minimum transition deadline of all new-mode tasks. Indeed, it follows from this statement and the definition of SM-MSO that all the transition deadlines would be met during every mode transition.

We provided a proof for the first key result in [31] (the proof is replicated in Section 5, page 17), and this result holds for any uniform platform (including identical platforms). About the second key result, it is worth noticing that there is *no job release* (and therefore no preemption) during every transition phase since we consider only FJP schedulers and all the old-mode tasks are disabled upon any mode change request. As a consequence, the length of any transition phase corresponds to *the time needed to complete all the rem-jobs* (this clearly appears in Figure 5). In the literature (and hereafter as well), the time needed to complete a given set of synchronous jobs upon a given platform is called the *makespan* defined as follows.

Definition 11 (Makespan) Let $J = \{J_1, J_2, \dots, J_n\}$ denote any set of n jobs of processing times c_1, c_2, \dots, c_n . Let π denotes any uniform multiprocessor platform composed of m CPUs. Let \mathcal{P} denote any job priority assignment and S denotes the schedule of J upon π by any work-conserving scheduler (including weakly and strongly work-conserving schedulers) using the priority assignment \mathcal{P} . The makespan denoted by $\text{ms}(J, \pi, \mathcal{P})$ is the earliest instant in S such that the n jobs of J are completed.

According to Definition 11, the length of any transition phase corresponds to the makespan generated by the set of jobs that are active in the system when the mode change is requested, i.e., the set of rem-jobs. Since the value of the makespan obviously depends on the number and processing times of the jobs (as well as on the CPU speeds), then the length of any transition phase from any mode M^i to any other mode M^j depends on both the number of rem-jobs and their remaining processing time at time $t_{\text{MCR}(j)}$. From this observation, determining an *upper-bound* on the makespan requires to consider the worst-case scenario, i.e., the scenario in which the number and the remaining processing time of the rem-jobs at time $t_{\text{MCR}(j)}$ is such that the generated makespan is maximum. This worst-case scenario is thus entirely defined by a specific set of rem-jobs that we name the *critical rem-job set* defined as follows.

Definition 12 (Critical rem-job set $\mathcal{J}_i^{\text{wc}}$) Assuming any transition from a specific mode M^i to any other mode M^j , the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ is the set of jobs issued from the tasks of τ^i that leads to the largest makespan.

For any work-conserving FJP scheduler (including FTP schedulers) and uniform platform (including identical platform), we will show that the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ of every transition from mode M^i to mode M^j is the one where each task τ_k^i has a rem-job at time $t_{\text{MCR}(j)}$ with a remaining processing time equals to C_k^i (i.e., the WCET of τ_k^i). This result is very intuitive: the makespan is as large as the number and processing times of the rem-jobs are large.

In this paper we address the problem of establishing mathematical expressions that provide the *maximum makespan* for any given set of *synchronous*⁷ jobs and especially for the critical rem-job set during each mode transition. This intention stems from the fact that the knowledge of the maximum makespan allows us to assert (or refute) that every new-mode task will meet its transition deadline during any mode transition using SM-MSO, thus ensuring the validity of SM-MSO for a given application τ and platform π as follows.

Validity Test 1 (For protocol SM-MSO) For any multi-mode real-time application τ and any uniform multiprocessor platform π , protocol SM-MSO is valid provided that, for every mode M^i ,

$$\overline{\text{ms}}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i) \leq \min_{j \neq i} \left\{ \min_{1 \leq k \leq n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\} \quad (2)$$

where \mathcal{P}^i is the job priority assignment derived from the old-mode scheduler \mathcal{S}^i and $\overline{\text{ms}}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i)$ is an upper-bound on the makespan, considering the set $\mathcal{J}_i^{\text{wc}}$ of jobs, the platform π and the job priority assignment \mathcal{P}^i .

⁷During every mode transition, the considered jobs are assumed to be synchronous because every rem-job is active and ready to execute upon the mode change request.

The above expression can be interpreted as follows: all the transition deadlines will be met during the execution of the system if, for every mode M^i , the maximum makespan (i.e., the maximum transition latency) generated by the rem-jobs issued from the tasks of τ^i cannot be larger than the minimum transition deadline of every task of every mode M^j .

This validity test is a *sufficient* condition that indicates, a priori, if all the deadlines will be met during all possible mode changes using the protocol SM-MSO. Unfortunately, to the best of our knowledge, the problem of determining the maximum makespan has never been studied in the literature. Rather, authors usually address the problem of determining a job priority assignment that minimizes the makespan [21, 19]. The goal in that framework being to ultimately reduce the completion times of the jobs as much as possible. This problem of finding priorities that minimize the makespan can be cast as a strongly NP-hard bin-packing problem [21, 19] for which numerous heuristics have been proposed in the literature. On the contrary, we provide in Sections 6–9 different *upper-bounds* on the makespan, assuming in turn *identical platforms and FJP schedulers*, *identical platforms and FTP schedulers*, *uniform platforms and FJP schedulers* and finally, *uniform platforms and FTP schedulers*.

3.3 FTP schedulers vs. FJP schedulers

As mentioned in Section 2.4, FTP schedulers are a particular case of FJP schedulers. However the remainder of this study distinguishes between these two scheduler families because FTP schedulers allow to determining a more precise upper-bound $\overline{ms}(\mathcal{J}_i^{wc}, \pi, \mathcal{P}^i)$ than FJP schedulers. The reason of this stems from the fact that the priority of each task (and thus the priority of every job) is known *at system design-time* for FTP schedulers whereas it is *unknown* beforehand for FJP schedulers.

At first blush, assuming that the job priority assignment \mathcal{P}^i is unknown for FJP schedulers can seem inconsistent since during every mode transition, we consider the critical rem-job set in the computation of $\overline{ms}(\mathcal{J}_i^{wc}, \pi, \mathcal{P}^i)$ (and this critical rem-job set is determined *at system design-time*). Therefore, it could be thought that \mathcal{P}^i can simply be derived from \mathcal{J}_i^{wc} . But this intuition is *erroneous* because for a given FJP scheduler, several job priority assignments can be derived from the same critical rem-job set as shown in the following example. Actually, given set of jobs, we are not aware of any job priority assignment leading to the maximum makespan.

Example 3 Let us consider a platform π composed of only 2 identical CPUs and an application τ composed of 2 modes M^i and M^j . Suppose that a mode change is requested from M^i to M^j and the old-mode scheduler S^i is EDF. The old-mode M^i contains 3 tasks with characteristics given in Table 2.

Tasks	C_k^i	$D_k^i = T_k^i$
τ_1^i	5	15
τ_2^i	5	16
τ_3^i	7	18

Table 2: Characteristics of the tasks in M^i .

As introduced earlier, the critical rem-job set for this mode transition is given by $\mathcal{J}_i^{wc} = \{J_1, J_2, J_3\}$ with processing time C_1^i, C_2^i and C_3^i , respectively. This will be formally proved in Corollary 1 (on page 19), assuming any FJP scheduler and any uniform platform. Actually, this critical rem-job set specifies only the processing time of the jobs, not the release time, neither the absolute deadline. Consequently, different job priority assignments can be derived from \mathcal{J}_i^{wc} . We depict two of them in Figures 6 and 7. In both figures the time is relative to the instant $t_{MCR(j)}$ (i.e., $t_{MCR(j)} = 0$). The release time and the absolute deadline of each job J_k are denoted by a_k and d_k , respectively. These two job priority assignments are obtained as follows.

Job priority assignment 1. If we assume that the three jobs are released exactly at the MCR invoking time $t_{MCR(j)}$, i.e., $a_1 = a_2 = a_3 = t_{MCR(j)}$, then the absolute deadline of each job J_k is given by $d_k \stackrel{\text{def}}{=} t_{MCR(j)} + D_k^i$. In Figure 6, the deadline of each job is thus: $d_1 = 15$, $d_2 = 16$ and $d_3 = 18$ and according to EDF, this leads to the job priority assignment $J_1 >_{\text{EDF}} J_2 >_{\text{EDF}} J_3$ (and to a makespan of 12).

Job priority assignment 2. Starting from the previous release pattern in which all the jobs are released simultaneously at time $t_{MCR(j)}$, one can slightly move backward the release time of job J_3 (for instance) in such a manner that J_3 is released at time $t_{MCR(j)} - 5$ (see Figure 7). Its absolute deadline d_3 is thus shifted to time $t_{MCR(j)} + 13$ and since no assumption is made about the schedule before time $t_{MCR(j)}$, we can suppose that J_3 did not execute before $t_{MCR(j)}$. Therefore, the processing time of J_3 at time $t_{MCR(j)}$ is $C_3^i = 5$ and the job priority assignment resulting from this new release pattern is $J_3 >_{\text{EDF}} J_2 >_{\text{EDF}} J_1$ (leading to a makespan of 10).

In the particular case of EDF, shifting the absolute deadline of these three jobs by distinct amplitudes can modify their relative priorities and a possibly large number of job priority assignments can be derived from the same critical rem-job set \mathcal{J}_i^{wc} .

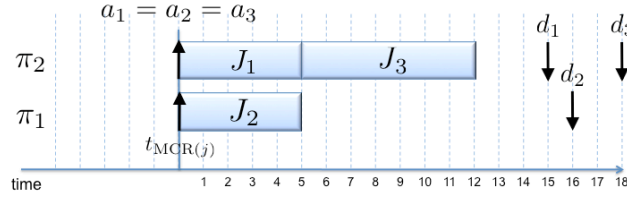


Figure 6: Assuming that the three jobs are released simultaneously upon the $MCR(j)$ allows to derive a first job priority assignment.

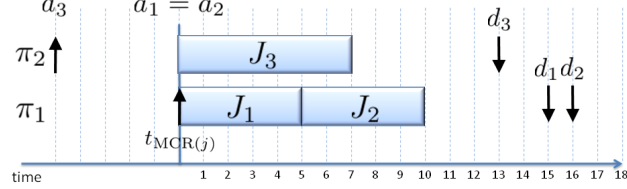


Figure 7: Another job priority assignment can be derived by slightly modifying the release pattern of the jobs. Note that this modification leads to another makespan.

Because the prior knowledge of the critical rem-job set does not allow determining a unique job priority assignment, FJP schedulers require to consider every possible job priority assignment in order to determine an upper-bound on the makespan. Hence, we refine the notation of $\overline{ms}(J, \pi, \mathcal{P}^i)$ as follows: the upper-bound on the makespan is denoted by $\overline{ms}(J, \pi, \mathcal{P})$ when \mathcal{P} is explicitly specified (in the context of FTP scheduler) and by $\overline{ms}(J, \pi)$ otherwise (in the context of FJP scheduler), with the interpretation that for every job priority assignment \mathcal{X} :

$$\overline{ms}(J, \pi) \geq ms(J, \pi, \mathcal{X})$$

It goes without saying that the prior knowledge of the jobs priority assignment allows for establishing tighter upper-bounds on the makespan, i.e., the upper-bound $\overline{ms}(J, \pi, \mathcal{P})$ is tighter than $\overline{ms}(J, \pi)$. From these refined notations, Expression 2 of Validity Test 1 can be rewritten as

$$\overline{ms}(\mathcal{J}_i^{wc}, \pi) \leq \min_{j \neq i} \left\{ \min_{1 \leq k \leq n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

for FJP schedulers, and as

$$\overline{ms}(\mathcal{J}_i^{wc}, \pi, \mathcal{P}^i) \leq \min_{j \neq i} \left\{ \min_{1 \leq k \leq n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

for FTP schedulers, where \mathcal{P}^i is the job priority assignment derived from the old-mode FTP scheduler \mathcal{S}^i .

4 The asynchronous protocol AM-MSO

4.1 Description of the protocol

The protocol AM-MSO (which stands for “Asynchronous Multiprocessor Minimum Single Offset” protocol) is an asynchronous version of the protocol SM-MSO. This protocol supports both uniform and identical platforms. The main idea of this second protocol is to reduce the delay applied to the enablement of the new-mode tasks, by enabling them as soon as possible. In contrast to SM-MSO, rem-jobs and new-mode tasks can be scheduled *simultaneously* during the transition phases according to the scheduler \mathcal{S}^{trans} defined as follows: (i) the priorities of the rem-jobs are assigned according to the old-mode scheduler; (ii) the priorities of the new-mode jobs are assigned according to the new-mode scheduler, and (iii) the priority of each rem-job is higher than the priority of every new-mode job.

Formally, suppose that the system is transitioning from mode M^{old} to mode M^{new} and let J_i and J_j be two active jobs during this transition. According to these notations we have $J_j >_{\mathcal{S}^{trans}} J_i$ if and only if one of the following conditions is satisfied:

$$\begin{aligned} & (J_j \in M^{old} \text{ and } J_i \in M^{new}) \\ \text{or } & (J_j \in M^{old} \text{ and } J_i \in M^{old} \text{ and } J_j >_{\mathcal{S}^{old}} J_i) \\ \text{or } & (J_j \in M^{new} \text{ and } J_i \in M^{new} \text{ and } J_j >_{\mathcal{S}^{new}} J_i) \end{aligned}$$

AM-MSO proceeds as follows: upon a $\text{MCR}(j)$, $\forall j$, all the old-mode tasks are disabled and the rem-jobs continue to be scheduled by S^i (assuming that M^i is the old-mode). Whenever any rem-job completes (say at time t), if there is no more waiting rem-jobs AM-MSO immediately enables some new-mode tasks, in contrast to SM-MSO which waits for the completion of *all* the rem-jobs. In order to select the new-mode tasks to enable at time t , AM-MSO uses the following *heuristic*: it considers every disabled new-mode task by non-decreasing order of transition deadline and enables those which can be scheduled by S^j upon the current available CPUs, i.e., the CPUs that are not running a rem-job and are therefore available for executing some new-mode tasks. The following example illustrates how AM-MSO manages mode transitions.

Example 4 Let us consider the same task sets as in Example 2. Figure 8 illustrates the AM-MSO transition protocol on a 2-processors platform.

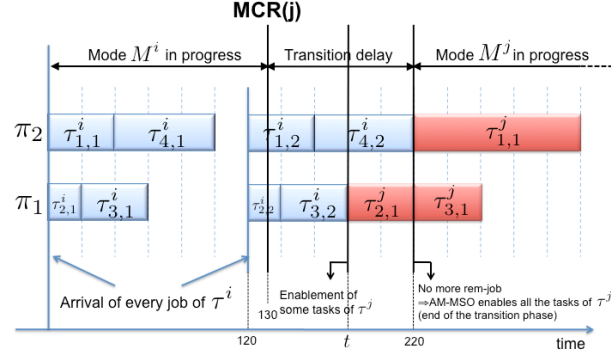


Figure 8: Illustration of a mode transition handled by AM-MSO.

Similarly to protocol SM-MSO, AM-MSO schedules the rem-jobs according to the old-mode scheduler from time $t_{\text{MCR}(j)} = 130$ to time t . Then at time t , the rem-job $\tau_{3,2}^j$ completes on CPU π_1 and there is no more waiting rem-jobs. Here AM-MSO reacts differently from SM-MSO: it scans every disabled task of τ^j (in non-decreasing order of transition deadline) and enables some of them in such a manner that the resulting set of enabled new-mode tasks can be scheduled by S^j upon 1 CPU (since at this time t , only the CPU π_1 is available for executing some new-mode tasks). We actually have no guarantee that scanning all the disabled tasks in non-decreasing order of transition deadline is optimal, but this heuristic appears as the most intuitive choice. At time 220, AM-MSO performs the same treatment as at time t . But since we assumed that every task set τ^k , $\forall k$, is schedulable by S^k on π , we know that all the remaining disabled new-mode tasks can be enabled at this time 220.

Notice that, in contrast to SM-MSO, the protocol AM-MSO allows mode changes to be requested during the mode transitions *only* until some new-mode tasks have been enabled (the instant t in Figure 8). Indeed, if the system is transitioning from any mode M^i to any other mode M^j and a mode change is requested to any mode M^z before time t , then AM-MSO can consider that the system is transitioning from mode M^i to mode M^z and the new-mode therefore becomes the mode M^z . However after time t , some tasks of mode M^j have already been enabled and AM-MSO does not allow the system to request any other mode change until the end of the transition phase from M^i to M^j , i.e., until all the tasks of mode M^j are enabled.

In order to determine whether a task can be safely enabled, protocol AM-MSO uses a binary function $\text{sched}(\pi, S, \tau^\ell)$ that returns True if and only if the task set τ^ℓ is schedulable by S upon π . This function is essential as we must always guarantee that all the deadlines are met for *all* the jobs in the system, including the deadlines of *all* the new-mode jobs. Considering a specific scheduler S , such a function can be derived from schedulability tests proposed for S in the literature⁸. Algorithm 9 provides a pseudo-code for protocol AM-MSO.

Observation 1 The whole “if-else-endif” block within lines 11–17 could be replaced with $\pi^{\text{avl}} \leftarrow \pi^{\text{avl}} \cup \{\pi_{m-r}\}$ as adding π_{m-r} (instead of π_ℓ) to π^{avl} does not make any difference if π is identical. However, we preferred to provide the reader with this longer version of the algorithm for sake of pedagogy. The shorter version explained here will be used in the Validity Algorithm 10 presented on page 15.

⁸To the best of our knowledge, there is no efficient *necessary and sufficient* schedulability test for any multiprocessor scheduler that complies with the requirements specified in Section 2.4. Theodore Baker has proposed in [6] a necessary and sufficient schedulability test for arbitrary-deadline sporadic tasks scheduled by Global-EDF but its time-complexity is very high so only small applications can be tested. Fortunately, many sufficient schedulability tests have been proposed for scheduler such as Global-EDF (see for instance [5, 7, 13, 10, 16]) and Global-DM (see for instance [4, 12, 11]).

Input: M^i : the old mode
Input: M^j : the new-mode
Input: the rem-jobs
Input: t : the current time during the transition
Input: π : the platform (uniform or identical)

```

1: if ( $t$  is the MCR invoking time) then
2:   Disable all the tasks of  $\tau^i$ 
3:   Sort the task set “disabled( $\tau^j, t$ )” by non-decreasing order of transition deadlines
4:    $\pi^{\text{avl}} \leftarrow \emptyset$ 
5: end if
6: Schedule the rem-jobs according to  $\mathcal{S}^{\text{trans}}$ 
7: if (any rem-job  $J_k$  completes at  $t$  on any CPU  $\pi_\ell$ ) then
8:    $r \leftarrow$  number of active rem-jobs at time  $t$ 
9:   if ( $r < m$ ) then
10:    /* Due to the completion of  $J_k$ , one CPU  $\notin \pi^{\text{avl}}$  becomes available. */
11:    if ( $\pi$  is identical) then
12:      /* The scheduler is weakly work-conserving. Thus, the CPU that becomes available is  $\pi_\ell$  */
13:       $\pi^{\text{avl}} \leftarrow \pi^{\text{avl}} \cup \{\pi_\ell\}$ 
14:    else
15:      /* The scheduler is strongly work-conserving. Thus, the CPU that becomes available is the
        ( $m - r$ )th slowest CPU. */
16:       $\pi^{\text{avl}} \leftarrow \pi^{\text{avl}} \cup \{\pi_{m-r}\}$ 
17:    end if
18:  end if
19:  for each  $\tau_r^j \in \text{disabled}(\tau^j, t)$  do
20:     $\tau^{\text{temp}} \leftarrow \text{enabled}(\tau^j, t) \cup \{\tau_r^j\}$ 
21:    if (sched( $\pi^{\text{avl}}, \mathcal{S}^j, \tau^{\text{temp}}$ )) then
22:      enable  $\tau_r^j$ 
23:    end if
24:  end for
25:  if ( $r = 0$ ) then
26:    enter the new-mode  $M^j$ 
27:  else
28:    Schedule all the rem-jobs and new-mode jobs according to  $\mathcal{S}^{\text{trans}}$ 
29:  end if
30: end if
  
```

Figure 9: AM-MSO protocol

4.2 Design of a validity test

For a given application τ and platform π , the main idea to determine whether AM-MSO allows to meet all the transition deadlines is to run Algorithm 9 for every possible mode transition, while considering the worst-case scenario for each one—the scenario in which the new-mode tasks are enabled as late as possible. From our definition of protocol AM-MSO, we know that every instant at which some new-mode tasks are enabled corresponds to an instant at which at least one CPU has no more rem-job to execute, i.e., an “idle-instant” defined as follows.

Definition 13 (Idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$) Let $J = \{J_1, J_2, \dots, J_n\}$ be any finite set of n synchronous jobs. Let π be a uniform multiprocessor platform and let \mathcal{P} be the job priority assignment used during the schedule of J upon π . If S denotes that schedule then the idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (with $k = 1, \dots, m$) is the earliest instant in S such that at least k CPUs idle.

By definition of the protocol AM-MSO, and in particular from the definition of $\mathcal{S}^{\text{trans}}$, a new-mode job never preempts a rem-job during the transition phases. Thereby, during every transition phase, new-mode tasks are enabled at each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ ($\forall k = 1, \dots, m$) where J is the set of rem-jobs at the MCR invoking time and \mathcal{P} is the job priority assignment derived from the old-mode scheduler when the mode change is requested. For obvious reasons, the *exact* values of these idle-instants depend on both the number of jobs in J and their *actual* execution times. Therefore, these exact value cannot be determined at system design-time and the main idea of our validity test is the following.

First, for every mode M^i we determine the set J of rem-jobs that leads to the largest idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ ($\forall k \in [1, m]$). From this point forward, we thus refine the definition of the *critical rem-job set* as follows.

Definition 14 (Critical rem-job set $\mathcal{J}_i^{\text{wc}}$) Assuming any transition from a specific mode M^i to any other mode M^j , the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ is the set of jobs issued from the tasks of τ^i that leads to the largest idle-instants.

As it will be shown in Corollary 1 (page 19), the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ of every mode M^i is the one that contains one job J_ℓ for each task τ_ℓ^i and such that every job $J_\ell \in \mathcal{J}_i^{\text{wc}}$ has a processing time equals to C_ℓ^i , i.e., the WCET of τ_ℓ^i . Informally speaking, the worst-case scenario during any mode transition is the one in which (i) every old-mode task releases a job *exactly* when the mode change is requested and (ii) every released job executes for its WCET.

Second, we determine (for any given set J of jobs) an upper-bound on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (for $k = 1, 2, \dots, m$). As in the previous section (and for the same reason), we distinguish between FTP and FJP schedulers. That is, for FTP schedulers we focus on determining an upper-bound $\text{idle}_k(J, \pi, \mathcal{P})$ on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (for $k = 1, 2, \dots, m$) assuming that the job priority assignment \mathcal{P} is known beforehand, whereas for FJP schedulers, we determine an upper-bound $\text{idle}_k(J, \pi)$ on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$, with the interpretation that for every job priority assignment \mathcal{A} :

$$\overline{\text{idle}}_k(J, \pi) \geq \text{idle}_k(J, \pi, \mathcal{A})$$

Finally, we simulate Algorithm 9 at each of these upper-bounds. That is, we verify whether all the transition deadlines are met while enabling the new-mode tasks at each instant $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P})$ (or $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ depending on the family of the old-mode scheduler). Obviously, if every transition deadline is met during this simulation then it will be met during the *actual* execution of the application.

It goes without saying that the prior knowledge of the jobs priority assignment allows for establishing tighter upper-bounds on the idle-instants, i.e., the upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ are tighter than $\overline{\text{idle}}_k(J, \pi)$. Notice that it results from these notations that $\overline{\text{idle}}_m(J, \pi)$ and $\overline{\text{idle}}_m(J, \pi, \mathcal{P})$ correspond to the upper-bounds $\overline{\text{ms}}(\mathcal{J}_i^{\text{wc}}, \pi)$ and $\overline{\text{ms}}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i)$ introduced in Validity Test 1, respectively.

Mathematical expressions of these upper-bounds $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ and $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i)$ on the k^{th} idle-instants are defined for both identical and uniform platforms in Sections 6–9. Algorithm 10 provides details on the validity test for AM-MSO, where the upper-bounds $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ must be replaced with $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i)$ at line 9 if the old-mode scheduler is FTP.

Notice that Algorithm 10 enables new-mode tasks only at the instants $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ (with $k = 1, 2, \dots, m$). That is, it implicitly considers that every instant at which CPUs become available to the new-mode tasks are as late as possible. As a consequence, if all the transition deadlines are met while running Algorithm 10 then all these deadlines will be met during every transition phase at run-time⁹. Nevertheless, the fact that Algorithm 10 simulates every idle-instant of every mode transition by its corresponding upper-bound $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ brings about the following situation: during the actual execution of the application, there could be some intervals of time (during any mode transition) during which the set of currently enabled new-mode tasks benefits from more (and *faster*) CPUs than during the execution of Algorithm 10. This kind of situation can occur upon identical and uniform platforms and for both FJP and FTP schedulers as shown in the following example.

⁹Because Algorithm 10 considers every transition between every pair of modes of the application.

Input: $\tau = \{\tau^1, \tau^2, \dots, \tau^x\}$

- 1: **for** (all $i, j \in [1, x]$ such that $i \neq j$) **do**
- 2: $\tau^{\text{disabled}} \leftarrow \tau^j$
- 3: $\tau^{\text{enabled}} \leftarrow \emptyset$
- 4: $\pi^{\text{avl}} \leftarrow \emptyset$
- 5: Sort τ^{disabled} by non-decreasing order of transition dead-
lines
- 6: **for** ($k = 1; k \leq m; k++$) **do**
- 7: $\pi^{\text{avl}} \leftarrow \pi^{\text{avl}} \cup \pi_k$
- 8: **for** (all $\tau_r^j \in \tau^{\text{disabled}}$) **do**
- 9: **if** ($\mathcal{D}_r^j(M^i) < \text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi)$) **then**
- 10: **return false**
- 11: **end if**
- 12: **if** ($\text{sched}(\pi^{\text{avl}}, \mathcal{S}^j, \tau^{\text{enabled}} \cup \{\tau_r^j\})$) **then**
- 13: $\tau^{\text{enabled}} \leftarrow \tau^{\text{enabled}} \cup \{\tau_r^j\}$
- 14: $\tau^{\text{disabled}} \leftarrow \tau^{\text{disabled}} \setminus \{\tau_r^j\}$
- 15: **end if**
- 16: **end for**
- 17: **end for**
- 18: **end for**
- 19: **return true**

Figure 10: Validity Test for AM-MSO

Example 5 Let us consider a 5-processors uniform platform π and a system which is transitioning from mode M^i to mode M^j . Other details such as the CPU speeds, the characteristics of the jobs and the job priority assignment are not relevant in the scope of this example. Figures 11 and 12 illustrate a situation where during some intervals of time the set of currently enabled new-mode tasks benefits from more (and faster) CPUs than during the execution of Algorithm 10.

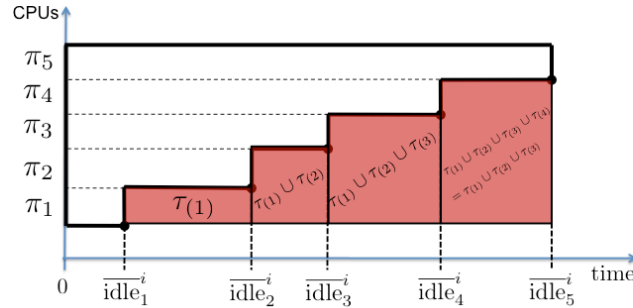


Figure 11: Illustration of the schedule assumed by the execution of Algorithm 10. In this schedule, new-mode tasks are enabled at each instant idle_k^i , $1 \leq k \leq m$.

For sake of clarity, Figure 12 uses the notations idle_k^i and $\overline{\text{idle}}_k^i$ instead of $\text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ and $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$, respectively. In this latter schedule, there can be less rem-jobs and/or rem-jobs with lower processing times than in the schedule of Figure 11 since the schedule of Figure 11 is drawn while assuming the critical rem-job set of mode M^i . This is the reason why the schedule of Figure 12 seems less “loaded” than the one of Figure 11. Due to the fact that (i) the validity test provided by Algorithm 10 uses the same function $\text{sched}(\pi, \mathcal{S}, \tau)$ as protocol AM-MSO at run-time and (ii) this function $\text{sched}(\pi, \mathcal{S}, \tau)$ is independent of the current time, we know that the set of tasks enabled at each instant idle_k^i ($k = 1, 2, \dots, m$) in Figure 12 is the same as the set of tasks enabled at each instant $\overline{\text{idle}}_k^i$ in Figure 11. Let us temporarily name this property the “equivalence property”. Let $\tau_{(k)}$ temporarily denote the set of tasks enabled at time idle_k^i , $\forall k \in [1, m]$ and suppose that at time idle_3^i in Figure 11 some tasks are enabled (i.e., $\tau_{(3)} \neq \phi$) and at time $\overline{\text{idle}}_4^i$ no task is enabled, i.e., $\tau_{(4)} = \phi$. Thanks to the equivalence property, we know that the tasks enabled at time idle_3^i in Figure 12 are the tasks of $\tau_{(3)}$ and those enabled at time $\overline{\text{idle}}_4^i$ are the tasks of $\tau_{(4)}$. Since we assumed in Figure 12 that $\text{idle}_3^i = \overline{\text{idle}}_4^i$, it holds that the tasks enabled at time idle_3^i are the tasks of $\tau_{(3)} \cup \tau_{(4)} = \tau_{(3)}$ (since $\tau_{(4)} = \phi$). It follows that in the time interval $[\text{idle}_3^i, \overline{\text{idle}}_4^i]$, only 3 CPUs are available to

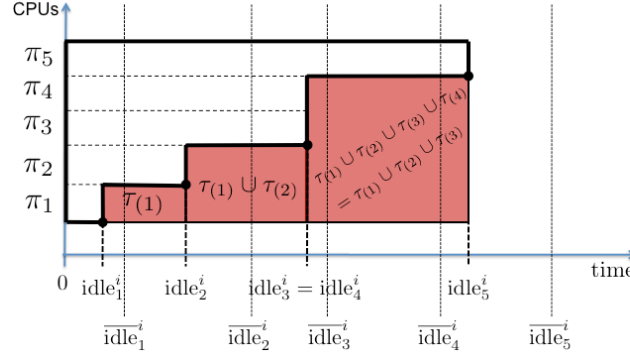


Figure 12: Illustration of a possible schedule during a transition from mode M^i to mode M^j in the *actual* execution of the application. Here, new-mode tasks are enabled at each instant idle_k , $1 \leq k \leq m$, where $\text{idle}_k \leq \bar{\text{idle}}_k$.

the task set $\tau_{(1)} \cup \tau_{(2)} \cup \tau_{(3)}$ in Figure 11 while 4 CPUs are available to this task set in Figure 12. Moreover, during this time interval, the additional CPU π_4 in Figure 12 is faster (or of equal speed) than every CPU in the subset of CPUs $\{\pi_1, \pi_2, \pi_3\}$ available to $\tau_{(1)} \cup \tau_{(2)} \cup \tau_{(3)}$ in Figure 11.

Lemma 3 proves that this kind of situation does not jeopardize the schedulability of the application during its execution.

Lemma 3 (See [27]) Any strongly work-conserving scheduler that is able to schedule a task set τ upon a uniform platform $\pi = [s_1, \dots, s_m]$ is also able to schedule τ upon any uniform platform π^* such that (i) $\pi^* \supseteq \pi$ and (ii) $\forall \pi_k \in \pi^*$ and $\pi_k \notin \pi$ we have $s_k \geq s_m$.

Proof 1 To obtain the proof, it is sufficient to show the lemma for $\pi^* = [s_1, \dots, s_m, s_{m+1}]$ where $s_{m+1} \geq s_m$. The proof is made by contradiction. Suppose there exists a task set τ that is schedulable by a strongly work-conserving scheduler S upon π , but not upon $\pi^* \supseteq \pi$. Consider the schedule upon π^* of a particular set J of jobs issued from τ that leads to a deadline miss, and let J^* be another set of jobs derived from J by reducing the processing time of each job J_i by the amount of time J_i executes upon the sub-platform $\pi^* \setminus \pi$, i.e., upon π_{m+1} . Since the scheduler is strongly work-conserving, the schedule of J by S upon the CPUs in common with π is the same as the one that would be produced by S for J^* upon platform π . Since a deadline is missed in the schedule of J upon π^* , then a deadline is missed also in the schedule of J^* upon π . But since the scheduler is predictable from Lemma 2, a deadline would be missed on π even (a fortiori) with the more demanding jobs set J , leading to a contradiction. The lemma follows.

Lemma 3 is proved while considering uniform platforms and strongly work-conserving schedulers but one can easily show that it also holds for identical platforms and weakly work-conserving schedulers.

5 Some basic results for determining validity tests

5.1 Introduction to the three required key results

Three key results are required to establish a validity test for SM-MSO and AM-MSO.

Key Result 1 It must be proved that disabling the old-mode tasks upon any MCR does not jeopardize the schedulability of the rem-jobs when they continue to be scheduled by the old-mode scheduler. That is, it must be guaranteed that the absolute deadline $d_{a,b}^i$ of every rem-job $\tau_{a,b}^i$ is met during any mode transition from every mode M^i .

Key Result 2 The critical rem-job set $\mathcal{J}_i^{\text{wc}}$ for every mode M^i must be determined. Indeed, for every mode transition from mode M^i to any other mode M^j , our validity test (see Algorithm 10) determines the upper-bounds on the idle-instants by basing the computations on the corresponding critical rem-job sets $\mathcal{J}_i^{\text{wc}}$ (at line 10). In all cases (i.e., identical or uniform platforms and FJP or FTP schedulers), we will provide a proof that the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ of every mode M^i is the one that contains one job J_ℓ for each task τ_ℓ^i and such that every job $J_\ell \in \mathcal{J}_i^{\text{wc}}$ has a processing time equals to C_ℓ^i , i.e., the WCET of the task τ_ℓ^i .

Key Result 3 A mathematical expression must be established that provides, for any given set J of jobs and platform π :

1. an upper-bound $\overline{\text{idle}}_k(J, \pi)$ ($1 \leq k \leq m$) on each idle-instant $\text{idle}_k(J, \pi, \mathcal{X})$, for every job priority assignment \mathcal{X} . This concerns FJP schedulers.
2. an upper-bound $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ ($1 \leq k \leq m$) on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$, for a specific job priority assignment \mathcal{P} . This concerns FTP schedulers.

Note that the protocol SM-MSO requires only an upper-bound on the makespan, i.e., on the m^{th} idle-instant $\overline{\text{idle}}_m(J, \pi)$ and $\overline{\text{idle}}_m(J, \pi, \mathcal{P})$.

5.2 Proof of the first key result

Lemma 4 proves the first key result introduced above for any uniform platform and strongly work-conserving scheduler, as well as any identical platform and weakly work-conserving scheduler. This result, which is essential to the validity tests of both protocols SM-MSO and AM-MSO, is based on the notion of *predictability* introduced on page 7. It has been drawn from [31] and extended to uniform platforms.

Lemma 4 *Let M^i and M^j denote two distinct modes of the application. If the application is running in mode M^i and a MCR(j) occurs at time $t_{\text{MCR}(j)}$ then every rem-job meets its deadline during the transition phase while being scheduled by the old-mode scheduler \mathcal{S}^i .*

Proof 2 *From our first assumption on page 7, the set of tasks τ^i of the mode M^i is schedulable by \mathcal{S}^i upon π . When the MCR(j) is invoked at time $t_{\text{MCR}(j)}$, the transition protocol disables every old-mode task, which is equivalent to set the processing time of all their future jobs to zero. Since \mathcal{S}^i is predictable (from Lemma 1 or 2 depending on the scheduler family), the deadline of every rem-job is still met in the produced schedule. The lemma follows.*

5.3 Proof of the second key result

Corollary 1 proves the second key result introduced above for any uniform platform and strongly work-conserving FTP (or FJP) scheduler, as well as any identical platform and weakly work-conserving FTP (or FJP) scheduler. It has been drawn from the following Lemma 5.

Lemma 5 *Let π be any uniform multiprocessor platforms (including identical platforms) and let J and J' be any fixed set of n synchronous jobs such that $J = \{J_1, J_2, \dots, J_n\}$ of processing times c_1, c_2, \dots, c_n and $J' = \{J'_1, J'_2, \dots, J'_n\}$ of processing times c'_1, c'_2, \dots, c'_n . For any job priority assignment \mathcal{P} , if there exists a bijective function between J and J' such that every job $J'_r \in J'$ is mapped to exactly one job $J_r \in J$ and such that $c'_r \leq c_r$, then the k^{th} idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ ($\forall k \in [1, m]$) in the schedule of J upon π is not lower than the k^{th} idle-instant $\text{idle}_k(J', \pi, \mathcal{P})$ in the schedule of J' , i.e., it holds $\forall k \in [1, m]$ that*

$$\text{idle}_k(J', \pi, \mathcal{P}) \leq \text{idle}_k(J, \pi, \mathcal{P})$$

Proof 3 *The proof is a consequence of the predictability of work-conserving schedulers (including both weakly and strongly work-conserving schedulers). Let S and S' denote the schedule of J and J' upon π with \mathcal{P} , respectively. We denote by comp_r and comp'_r the completion time of any job J_r in S and J'_r in S' , respectively. It follows from the fact that $c'_r \leq c_r$ ($\forall r \in [1, n]$) and from the predictability of the considered schedulers (see Lemma 1 or 2 depending on the scheduler family) that $\forall r \in [1, n]$:*

$$\text{comp}'_r \leq \text{comp}_r \tag{3}$$

The proof is made by contradiction. Suppose that there exists $\ell \in [1, m]$ such that

$$\text{idle}_\ell(J, \pi, \mathcal{P}) < \text{idle}_\ell(J', \pi, \mathcal{P})$$

Figures 13 and 14 illustrate an example of schedules S and S' on a 5-processors uniform platform, respectively, where $\text{idle}_3(J, \pi, \mathcal{P}) < \text{idle}_3(J', \pi, \mathcal{P})$.

Since the platform is uniform in these examples, the scheduler is strongly work-conserving and both schedules S and S' form a staircase. In both Figures 13 and 14, we voluntarily omit the details about the CPU speeds, the jobs characteristics, etc. since they are useless in the scope of these examples.

Similarly, Figures 15 and 16 illustrate an example of schedules S and S' on a 5-processors identical platform, respectively, where $\text{idle}_3(J, \pi, \mathcal{P}) < \text{idle}_3(J', \pi, \mathcal{P})$. Since the platform is identical in these examples, the scheduler is assumed to be weakly work-conserving. Furthermore, note that in both examples no job is released after time 0.

By definition of the idle-instants, the schedule of any set \mathcal{J} of jobs upon any uniform or identical multiprocessor platform is such that $\forall k \in [1, m]$:

- the idle-instant $\text{idle}_k(\mathcal{J}, \pi, \mathcal{P})$ corresponds to the completion time of a job,

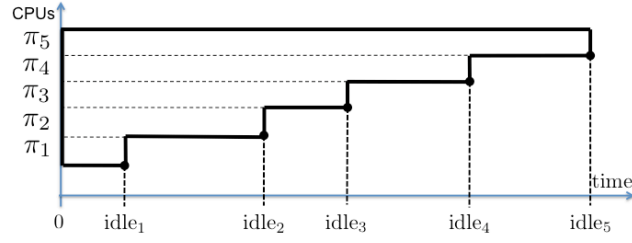


Figure 13: An example of schedule S upon a 5-processors uniform platform. The idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ are denoted by idle_k for sake of clarity.

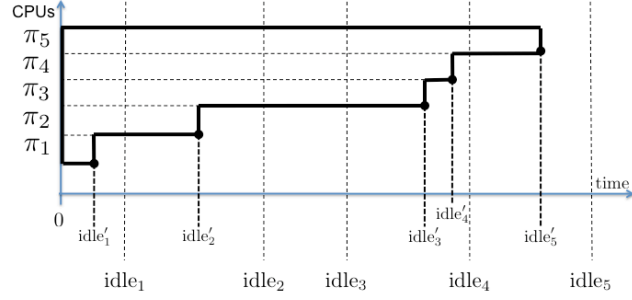


Figure 14: An example of schedule S' upon the same 5-processors uniform platform. Also for sake of clarity, the idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ and $\text{idle}_k(J', \pi, \mathcal{P})$ are denoted by idle_k and idle'_k , respectively. In this figure, we have by contradiction $\text{idle}_3 < \text{idle}'_3$.

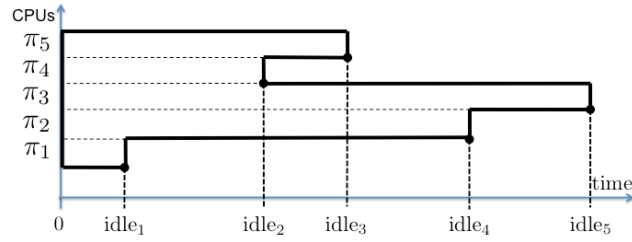


Figure 15: An example of schedule S upon a 5-processors identical platform. The idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ are denoted by idle_k for sake of clarity.

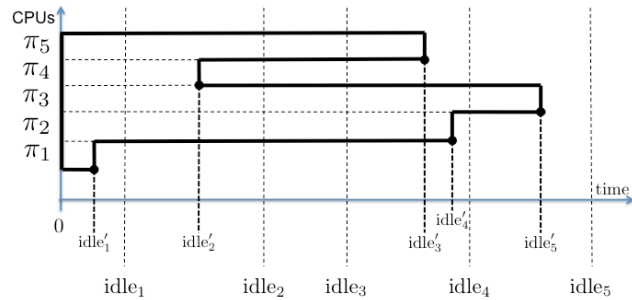


Figure 16: An example of schedule S' upon the same 5-processors identical platform. Also for sake of clarity, the idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ and $\text{idle}_k(J', \pi, \mathcal{P})$ are denoted by idle_k and idle'_k , respectively. In this figure, we have by contradiction $\text{idle}_3 < \text{idle}'_3$.

- *there is no waiting job at time $\text{idle}_k(J, \pi, \mathcal{P})$ and,*
- *there are at most $(m - k)$ running jobs at time $\text{idle}_k(J, \pi, \mathcal{P})$. “At most” since there can exist some $r > k$ such that $\text{idle}_r(J, \pi, \mathcal{P}) = \text{idle}_k(J, \pi, \mathcal{P})$.*

Since every idle-instant corresponds to the completion of a job, this implies that within the time interval $[\text{idle}_\ell(J, \pi, \mathcal{P}), \text{idle}_\ell(J', \pi, \mathcal{P})]$ there are at most $(m - \ell)$ running jobs in S while there are at least $(m - \ell + 1)$ running jobs in S' . Therefore, within $[\text{idle}_\ell(J, \pi, \mathcal{P}), \text{idle}_\ell(J', \pi, \mathcal{P})]$, at least one job (say J_r) is already completed in S while J'_r is still running in S' . The fact that J'_r completes later in S' than J_r in S leads to a direct contradiction of Inequality 3. As we can see in Figures 14 and 16, three jobs are running in S' during the time interval $[\text{idle}_3(J, \pi, \mathcal{P}), \text{idle}_3(J', \pi, \mathcal{P})]$ while only two jobs are running in S , meaning that there is one job which is completed in S and still running in S' . The lemma follows.

Corollary 1 For any uniform multiprocessor platforms π and for any transition of the system from mode M^i to mode M^j , let J^{any} denote any set of rem-jobs issued from the old-mode tasks and let $\mathcal{J}_i^{\text{wc}}$ be the set of rem-jobs that contains one job J_ℓ for each task τ_ℓ^i and such that every job $J_\ell \in \mathcal{J}_i^{\text{wc}}$ has a processing time equals to C_ℓ^i . The k^{th} idle-instants $\text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ ($\forall k \in [1, m]$) in the schedule of $\mathcal{J}_i^{\text{wc}}$ is never lower than the k^{th} idle-instant $\text{idle}_k(J^{\text{any}}, \pi)$ in the schedule of J^{any} , i.e., it holds $\forall k \in [1, m]$ that

$$\text{idle}_k(J^{\text{any}}, \pi) \leq \text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi)$$

Proof 4 The proof is a consequence of Lemma 5. Let c_r^{wc} and c_r^{any} denote the processing time of job J_r in $\mathcal{J}_i^{\text{wc}}$ and J^{any} , respectively. By definition, $\mathcal{J}_i^{\text{wc}}$ contains one job J_r of processing time C_r^i for each task $\tau_r^i \in \tau^i$, i.e., it holds $\forall \tau_r^i \in \tau^i$ that

$$c_r^{\text{wc}} = C_r^i$$

and thus we know by definition of J^{any} that $\forall J_r \in J^{\text{any}}$,

$$c_r^{\text{any}} \leq c_r^{\text{wc}}$$

In addition, we know that there could be some jobs $J_\ell \in \mathcal{J}_i^{\text{wc}}$ such that $J_\ell \notin J^{\text{any}}$ (since J^{any} does not necessarily contain one job for each old-mode task). For each such job J_ℓ we add a fake job J'_ℓ in J^{any} with $c_\ell^{\text{any}} = 0$. It results from this operation that the number of jobs in both $\mathcal{J}_i^{\text{wc}}$ and J^{any} are the same (we denote this number by n) and there is a bijective function between $\mathcal{J}_i^{\text{wc}}$ and J^{any} such that every job $J_r \in \mathcal{J}_i^{\text{wc}}$ is mapped to by exactly one job $J_r \in J^{\text{any}}$ and such that $c_r^{\text{any}} \leq c_r^{\text{wc}}$. Thanks to this bijection, we know from Lemma 5 that $\forall r \in [1, m]$ we have

$$\text{idle}_r(J^{\text{any}}, \pi) \leq \text{idle}_r(\mathcal{J}_i^{\text{wc}}, \pi)$$

and the corollary follows.

By definition, for every mode transition from any mode M^i upon π , each $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ is an upper-bound on the k^{th} idle-instant in the schedule of $\mathcal{J}_i^{\text{wc}}$ (this also holds for each upper-bound $\text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P})$ if the job priority assignment \mathcal{P} is known beforehand). Thanks to Corollary 1, we are now aware that each upper-bound $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ (and $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P})$) is also an upper-bound on the k^{th} idle-instant in the schedule of any other set of rem-jobs issued from the old-mode tasks (i.e., the tasks of τ^i). That is, for every mode transition from any mode M^i we have $\forall k \in [1, m]$: $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi) \geq \text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi) \geq \text{idle}_k(J^{\text{any}}, \pi)$ and $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}) \geq \text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}) \geq \text{idle}_k(J^{\text{any}}, \pi, \mathcal{P})$, where J^{any} denotes any set of rem-jobs issued from the tasks of τ^i . As a result, the instants $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ (and $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P})$), with $k = 1, 2, \dots, m$, can be considered as the *largest* instants at which new-mode tasks are enabled during every transition from mode M^i and thus, these instants can be used in our validity test given by Algorithm 10.

5.4 Organization for the third key result

The third key result consists in determining a mathematical expression for each upper-bound $\overline{\text{idle}}_k(J, \pi)$ (or $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$) depending on the scheduler family, i.e., FJP or FTP), for all $1 \leq k \leq m$. Depending on the type of the platform (uniform or identical) and on the scheduler family (FJP or FTP), we distinguish between four different cases that are studied in turn in the following four sections. More precisely:

- ▷ Section 6 addresses the *identical* and *FJP* case.
- ▷ Section 7 addresses the *identical* and *FTP* case.
- ▷ Section 8 addresses the *uniform* and *FJP* case.
- ▷ Section 9 addresses the *uniform* and *FTP* case.

Recall that the protocol SM-MSO requires only an upper-bound on the makespan, i.e., on the m^{th} idle-instant. The organization for the third key result is as follows.

6 Identical platforms and FJP schedulers

This section is organized as follows. First, Section 6.1 determines an *upper-bound* $\overline{\text{idle}}_k(J, \pi)$ on the earliest time-instant where at least k CPUs are idle and derives an upper-bound $\overline{\text{ms}}(J, \pi)$ on the maximum makespan. Then, Section 6.2 shows that this upper-bound $\overline{\text{ms}}(J, \pi)$ is 2-competitive, with the interpretation that $\overline{\text{ms}}(J, \pi)$ is at most twice the exact value of the maximum makespan. Finally, Section 6.3 establishes a *sufficient* validity test for protocols SM-MSO and AM-MSO.

6.1 Upper-bounds $\overline{\text{idle}}_k(J, \pi)$ on the idle-instants

Throughout this section, J refers to any set of n jobs. For sake of clarity, we will use the notation $\overline{\text{idle}}_k$ instead of $\overline{\text{idle}}_k(J, \pi)$ and similarly, we will use the notation idle_k to denote the *exact* value of the k^{th} idle-instant. Before introducing the computation of these upper-bounds $\overline{\text{idle}}_k$, $1 \leq k \leq m$, let us introduce the following result taken from [31].

Lemma 6 (See [31]) *Suppose that J is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \dots \leq c_n$. Then, whatever the job priority assignment we have $\forall j, k \in [1, m]$ such that $j < k$:*

$$\text{idle}_j \geq \text{idle}_k - c_{n-m+k}$$

Based on this Lemma 6, the following result was proved in our previous work [31].

Lemma 7 (See [31]) *Suppose that J is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \dots \leq c_n$. Then, whatever the job priority assignment, an upper-bound $\overline{\text{idle}}_k$ on the idle-instant idle_k , $1 \leq k \leq m$, is given by c_k if $n = m$ or by*

$$\max_{i=0}^{n-m+k-1} \left\{ \frac{\sum_{j=1}^n c_j - \sum_{j=i+1}^{i+m-k+1} c_j}{m} + \frac{\sum_{j=i+1}^{i+m-k+1} c_j}{m-k+1} \right\} \quad (4)$$

otherwise ($n > m$).

Holding this result, we improve here this previous analysis by (i) successfully establishing *another* upper-bound $\overline{\text{idle}}_k(J, \pi)$ on each idle-instant $\text{idle}_k(J, \pi)$ and (ii) proving that these alternative upper-bounds are always tighter than those proposed in Lemma 7. In short, we complete our previous work [31] as follows.

- ▷ Lemma 8 shows that Expression 4 of $\overline{\text{idle}}_k$ is always maximal for $i = n - m + k - 1$.
- ▷ Lemma 9 proposes another upper-bound $\overline{\text{idle}}_k(J, \pi)$ on each idle-instant idle_k .
- ▷ Lemma 10 shows that these alternative upper-bounds $\overline{\text{idle}}_k(J, \pi)$, $\forall k \in [1, m]$, are never larger than those provided by Expression 4.
- ▷ Finally, based on these alternative upper-bounds, Corollary 2 derives an upper-bound on the makespan.

Lemma 8 *If $n > m$, Expression 4 is maximal for $i = n - m + k - 1$.*

Proof 5 *This result is presented in Lemma 2.10 in [28]. Due to the space limitation and because the proof is simply based on algebra, we do not repeat it here.*

Thanks to Lemma 8, Expression 4 can be rewritten as follows: $\overline{\text{idle}}_k \stackrel{\text{def}}{=} c_k$ if $n = m$ or

$$\overline{\text{idle}}_k \stackrel{\text{def}}{=} \frac{\sum_{j=1}^n c_j - \sum_{j=n-m+k}^n c_j}{m} + \frac{\sum_{j=n-m+k}^n c_j}{m-k+1} \quad (5)$$

otherwise ($n > m$).

Lemma 9 *Suppose that J is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \dots \leq c_n$. Then, whatever the job priority assignment, an upper-bound $\overline{\text{idle}}_k$ on the idle-instant idle_k , $1 \leq k \leq m$, is given by $\overline{\text{idle}}_k \stackrel{\text{def}}{=} c_k$ if $n = m$ or by*

$$\overline{\text{idle}}_k \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n c_i + (k-1) \cdot c_{n-m+k}}{m} \quad (6)$$

otherwise ($n > m$).

Proof 6 *The case where $n = m$ is obvious. Otherwise, the proof is made by contradiction. Suppose that there exists $k \in [1, m]$ such that $\text{idle}_k > \overline{\text{idle}}_k$. The following properties hold:*

- **Prop. (a):** $\forall j > k: \text{idle}_j \geq \text{idle}_k$ (by definition of the idle-instants).
- **Prop. (b):** $\forall j < k: \text{idle}_j \geq \text{idle}_k - c_{n-m+k}$ (from Lemma 6).

The proof starts with this obvious equality:

$$\sum_{j=1}^m \text{idle}_j = \sum_{j=1}^{k-1} \text{idle}_j + \text{idle}_k + \sum_{j=k+1}^m \text{idle}_j$$

Then, applying properties (a) and (b) to the right-hand side yields

$$\begin{aligned} \sum_{j=1}^m \text{idle}_j &\geq \sum_{j=1}^{k-1} (\text{idle}_k - c_{n-m+k}) + \text{idle}_k + \sum_{j=k+1}^m \text{idle}_k \\ &\geq (k-1)(\text{idle}_k - c_{n-m+k}) + \text{idle}_k \\ &\quad + (m-k) \cdot \text{idle}_k \\ &\geq m \cdot \text{idle}_k - (k-1) \cdot c_{n-m+k} \end{aligned}$$

Since by hypothesis $\text{idle}_k > \overline{\text{idle}}_k$, replacing idle_k with $\overline{\text{idle}}_k$ in the above inequality leads to

$$\begin{aligned} \sum_{j=1}^m \text{idle}_j &> m \cdot \overline{\text{idle}}_k - (k-1) \cdot c_{n-m+k} \\ &> m \left(\frac{\sum_{i=1}^n c_i + (k-1) \cdot c_{n-m+k}}{m} \right) \\ &\quad - (k-1) \cdot c_{n-m+k} \\ &> \sum_{i=1}^n c_i \end{aligned}$$

This leads to a contradiction since it obviously holds by definition of the idle-instants that $\sum_{j=1}^m \text{idle}_j = \sum_{i=1}^n c_i$. The lemma follows.

Lemma 10 The upper-bounds $\overline{\text{idle}}_k$ (with $k = 1, 2, \dots, m$) provided by Expression 6 are never larger than those provided by Expression 5.

Proof 7 The proof is made by contradiction. Let k be any integer in $[1, m]$. Let $\overline{\text{idle}}_k^{\text{old}}$ and $\overline{\text{idle}}_k^{\text{new}}$ denote the upper-bound provided by Expressions 5 and 6, respectively, and suppose that $\overline{\text{idle}}_k^{\text{new}} > \overline{\text{idle}}_k^{\text{old}}$. From Expressions 5 and 6 we get

$$\begin{aligned} &\frac{\sum_{j=1}^n c_j + (k-1) \cdot c_{n-m+k}}{m} \\ &> \frac{\sum_{j=1}^n c_j - \sum_{j=n-m+k}^n c_j}{m} + \frac{\sum_{j=n-m+k}^n c_j}{m-k+1} \end{aligned}$$

By multiplying both sides by $m \cdot (m-k+1)$ we get

$$\begin{aligned} &(m-k+1) \cdot \left(\sum_{j=1}^n c_j + (k-1) \cdot c_{n-m+k} \right) \\ &> (m-k+1) \cdot \left(\sum_{j=1}^n c_j - \sum_{j=n-m+k}^n c_j \right) + m \sum_{j=n-m+k}^n c_j \end{aligned}$$

Thus,

$$(m-k+1) \cdot (k-1) \cdot c_{n-m+k} > (k-1) \cdot \sum_{j=n-m+k}^n c_j$$

If $k = 1$ then we obviously get $0 > 0$ and the lemma follows. Otherwise, if $k > 1$ then dividing both sides by $(k-1)$ yields

$$(m-k+1) \cdot c_{n-m+k} > \sum_{j=n-m+k}^n c_j$$

In this case, in the right-hand side of the above inequality, there are $m-k+1$ terms that are not lower than c_{n-m+k} each. This therefore leads to a contradiction since $c_1 \leq c_2 \leq \dots \leq c_n$. The lemma follows.

The following corollary derives an upper-bound on the makespan from $\overline{\text{idle}}_m$ provided by Expression 6.

Corollary 2 Suppose that J is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \dots \leq c_n$. Then, whatever the job priority assignment, an upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ on the makespan is given by $\overline{\text{ms}}^{\text{ident}}(J, \pi) \stackrel{\text{def}}{=} c_n$ if $n = m$, or by

$$\overline{\text{ms}}^{\text{ident}}(J, \pi) \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{n-1} c_i}{m} + c_n \quad (7)$$

otherwise.

Proof 8 Since the makespan corresponds to the m^{th} idle-instant, an upper-bound on the makespan is given by $\overline{\text{idle}}_m$. Therefore, the proof is obtained by simply replacing k with m in Expression 6.

6.2 Accuracy of the upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi)$

In this section, Lemma 11 proves that the upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ is 2-competitive, according to the following definition.

Definition 15 (α -competitive) Any upper-bound is said to be α -competitive if it provides at most α times the exact value of the approximated parameter.

This is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET. Without this assumption, the minimum makespan that could be produced is always 0 since it can always be the case that no old-mode task has an active job when the mode change is requested. For instance in Figure 5, the makespan would be zero if the $\text{MCR}(j)$ was released at time 110. However, in order to guarantee that our approach always provides an *upper-bound* on the makespan we have to consider the worst-case scenario in which every old-mode task releases a job *exactly* upon the mode change request and all these jobs executes for their WCET during the transition.

Lemma 11 For any set J of jobs sorted by non-decreasing job processing time and for any identical multiprocessor platform π composed of m CPUs, the upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ is 2-competitive.

Proof 9 Recall from Expression 7 that,

$$\overline{\text{ms}}^{\text{ident}}(J, \pi) \stackrel{\text{def}}{=} \begin{cases} c_n & \text{if } (n \leq m) \\ \frac{\sum_{i=1}^{n-1} c_i}{m} + c_n & \text{otherwise} \end{cases}$$

Let $\text{ms}(J, m)$ denote the exact makespan for the set J of jobs and the m identical CPUs. Since we do not have any mathematical expression for determining this exact makespan $\text{ms}(J, m)$, our analysis is performed while considering a lower-bound $\underline{\text{ms}}^{\text{ident}}(J, m)$ on the makespan rather than its exact value, i.e., α is determined in such a manner that

$$\frac{\overline{\text{ms}}^{\text{ident}}(J, \pi)}{\underline{\text{ms}}^{\text{ident}}(J, m)} \leq \alpha$$

where

$$\underline{\text{ms}}^{\text{ident}}(J, m) \stackrel{\text{def}}{=} \begin{cases} c_n & \text{if } n \leq m \\ \max \left\{ c_n, \frac{\sum_{i=1}^n c_i}{m} \right\} & \text{if } n > m \end{cases}$$

The case where $n \leq m$ obviously leads to $\alpha = 1$ since both $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ and $\underline{\text{ms}}^{\text{ident}}(J, \pi)$ return a makespan of c_n . Otherwise (if $n > m$) the “max” operator in the definition of $\underline{\text{ms}}^{\text{ident}}(J, m)$ leads to two different cases.

Case 1: If $c_n \geq \frac{\sum_{i=1}^n c_i}{m}$ then we get

$$\begin{aligned} \frac{\overline{\text{ms}}^{\text{ident}}(J, \pi)}{\underline{\text{ms}}^{\text{ident}}(J, m)} &\leq \frac{\frac{\sum_{i=1}^{n-1} c_i}{m} + c_n}{c_n} \\ &\leq \frac{\frac{\sum_{i=1}^n c_i}{m} + c_n}{c_n} \end{aligned}$$

and since in this case we have $c_n \geq \frac{\sum_{i=1}^n c_i}{m}$, it holds that

$$\begin{aligned} \frac{\overline{\text{ms}}^{\text{ident}}(J, \pi)}{\underline{\text{ms}}^{\text{ident}}(J, m)} &\leq \frac{c_n + c_n}{c_n} \\ &\leq 2 \end{aligned}$$

Case 2: If $c_n < \frac{\sum_{i=1}^n c_i}{m}$ then

$$\frac{\overline{\text{ms}}^{\text{ident}}(J, \pi)}{\underline{\text{ms}}^{\text{ident}}(J, m)} \leq \frac{\frac{\sum_{i=1}^{n-1} c_i}{m} + c_n}{\frac{\sum_{i=1}^n c_i}{m}}$$

and since in this case we have $c_n < \frac{\sum_{i=1}^n c_i}{m}$, it holds that

$$\begin{aligned} \frac{\overline{\text{ms}}^{\text{ident}}(J, \pi)}{\text{ms}(J, m)} &\leq \frac{\frac{\sum_{i=1}^n c_i}{m} + \frac{\sum_{i=1}^n c_i}{m}}{\frac{\sum_{i=1}^n c_i}{m}} \\ &\leq 2 \end{aligned}$$

The lemma follows.

It holds from Lemma 11 that, for any set J of jobs and any identical platform composed of m CPUs, the upper-bound on the maximum makespan provided by $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ is at most twice the exact value of the maximum makespan. Additionally we can show that in some particular cases as the one provided in the following example, the upper-bounds $\overline{\text{idle}}_k(J, \pi)$ ($\forall k \in [1, m]$) defined on page 20 are exact.

Example 6 Let us consider the set of 12 jobs with characteristics given in Table 3 to be scheduled on a 3-processors identical platform.

c_1	c_2	c_3	c_4	c_5	c_6
1	1	1	1	1	1
c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
3	3	6	6	9	12

Table 3: Processing times of the 12 jobs in J .

For this set of jobs,

- the upper-bound $\overline{\text{idle}}_1 = 15$ is reached with the job priority assignment $J_7 > J_9 > J_{10} > J_{12} > J_{11} > J_8 > J_1 > J_2 > J_3 > J_4 > J_5 > J_6$.
- the upper-bound $\overline{\text{idle}}_2 = 18$ is reached with the job priority assignment $J_{10} > J_9 > J_1 > J_2 > J_3 > J_4 > J_5 > J_6 > J_{12} > J_7 > J_8 > J_{11}$.
- the upper-bound $\overline{\text{idle}}_3 = 23$ is reached with the job priority assignment $J_7 > J_{11} > J_{10} > J_1 > J_2 > J_9 > J_8 > J_3 > J_5 > J_4 > J_6 > J_{12}$.

Due to the space limitation, we did not draw the schedules corresponding to these priority assignments.

6.3 Validity tests for SM-MSO and AM-MSO

From Corollaries 1 and 2, the sufficient validity test given by Test 1 on page 9 can be rewritten as follows.

Validity Test 2 (SM-MSO, Identical and FJP) For any multi-mode real-time application τ and any identical platform π composed of m CPUs, the protocol SM-MSO is valid provided that, for every mode M^i ,

$$\overline{\text{ms}}^{\text{ident}}(\mathcal{J}_i^{\text{wc}}, \pi) \leq \min_{j \neq i} \left\{ \min_{k=1}^{n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

where $\overline{\text{ms}}^{\text{ident}}(\mathcal{J}_i^{\text{wc}}, \pi)$ is defined as in Expression 7 and $\mathcal{J}_i^{\text{wc}}$ is defined as follows:

- ▷ $\mathcal{J}_i^{\text{wc}} \stackrel{\text{def}}{=} \{J_1, J_2, \dots, J_{n_i}\}$
- ▷ each job $J_k \in \mathcal{J}_i^{\text{wc}}$ has a processing time equal to the WCET C_k^i of task τ_k^i
- ▷ $\mathcal{J}_i^{\text{wc}}$ is sorted by non-decreasing processing time.

Concerning the protocol AM-MSO, the upper-bounds $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ (for all $1 \leq k \leq m$) defined as in Lemma 9 can be used at line 10 of the validity algorithm given by Algorithm 10 (on page 15).

7 Identical platforms and FTP schedulers

This section is organized as follows. First, Section 7.1 determines an upper-bound $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ for any given job priority assignment \mathcal{P} and derives an upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi, \mathcal{P})$ on the maximum makespan. Then, Section 7.2 shows that this upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ is 1-competitive, with the interpretation that $\overline{\text{ms}}^{\text{ident}}(J, \pi)$ corresponds to the exact value of the maximum makespan. Finally, Section 7.3 establishes a *sufficient* validity test for the protocols SM-MSO and AM-MSO.

7.1 Upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ on the idle-instants

As introduced earlier, this section focuses on determining a mathematical expression for the upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ where J refers to any set of n jobs, π denotes any identical multiprocessor platform composed of m CPUs and \mathcal{P} is a specific given job priority assignment. Indeed, for a given FTP scheduler the priority of every task (and thus of every job) is known beforehand. This prior knowledge allows us to determine tighter upper-bounds than those proposed in the previous section. Once again, for sake of clarity, we will use the notations idle_k and $\overline{\text{idle}}_k$ instead of $\text{idle}_k(J, \pi, \mathcal{P})$ and $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$, respectively.

For any transition from a given mode M^i to any other mode M^j , the knowledge of the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ and the fact that the job priority assignment is known beforehand allow us to compute the *exact* maximum idle-instants $\overline{\text{idle}}_k$ —exact in the sense that they are actually reached if every job executes for its WCET—simply by drawing the schedule of $\mathcal{J}_i^{\text{wc}}$ and by measuring the idle-instants idle_k in that schedule. Indeed, from Corollary 1 (on page 19), each idle-instant $\text{idle}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P})$ is an upper-bound on the idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ derived from the schedule of any other set J of rem-jobs. Before expressing these exact maximum idle-instants, let us introduce the following definition.

Definition 16 (Processed work Work_k^i) Let π denote any identical multiprocessor platform and let \mathcal{S} be any global, weakly work-conserving and FTP scheduler. Let $J = \{J_1, J_2, \dots, J_n\}$ denote any set of n jobs sorted by decreasing \mathcal{S} -priority, i.e., $J_1 >_{\mathcal{S}} J_2 >_{\mathcal{S}} \dots >_{\mathcal{S}} J_n$ and let S^i denote the schedule by \mathcal{S} of the i highest priority jobs of J upon π . The processed work Work_k^i ($1 \leq k \leq m$ and $0 \leq i \leq n$) denotes the amount of processing time executed on CPU π_k in S^i .

In order to familiarize the reader with this notation Work_k^i , we provide the following example.

Example 7 Let us consider the set J of 7 jobs with characteristics given in Table 4 to be scheduled on a 4-processors identical platform, following the priority assignment: $J_1 > J_2 > \dots > J_7$.

c_1	c_2	c_3	c_4	c_5	c_6	c_7
7	2	5	16	6	5	5

Table 4: Processing times of the 7 jobs in J .

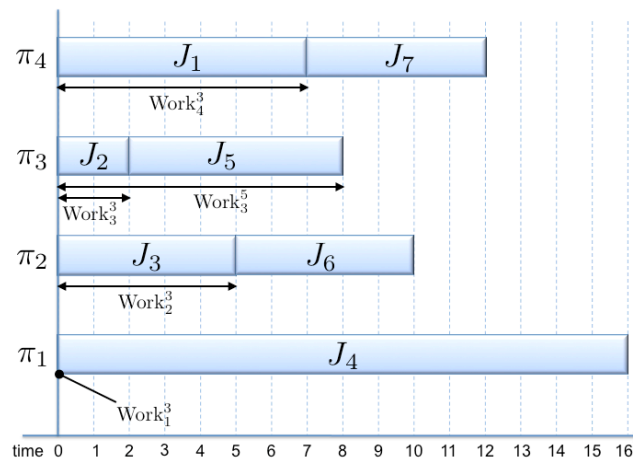


Figure 17: Illustration of the notion of processed work Work_k^i .

Figure 17 illustrates the schedule of J upon the 4 CPUs. In this schedule, we have $\text{Work}_3^5 = 8$ because, in the schedule S^5 of the 5 highest priority jobs J_1, J_2, J_3, J_4, J_5 , the amount of processing time units executed on π_3 is

$c_2 + c_5 = 8$. Similarly, $\text{Work}_4^3 = 7, \text{Work}_3^3 = 2, \text{Work}_2^3 = 5$ and $\text{Work}_1^3 = 0$ because, in the schedule S^3 of jobs J_1, J_2, J_3 , we can see that 7 processing time units are executed on π_4 (i.e., job J_1), 2 processing time units are executed on π_3 (i.e., job J_2), 5 processing time units are executed on π_2 (i.e., job J_3) and no processing time unit is executed on π_1 . Notice that $\text{Work}_k^0 = 0 \forall k = 1, 2, \dots, m$.

Lemma 12 provides the exact values of Work_k^i ($\forall i \in [1, n]$ and $\forall k \in [1, m]$) when each job executes for its WCET. Then, Corollary 3 derives the exact maximum idle-instants idle_k $1 \leq k \leq m$, for the scheduling of any set J of n jobs upon any m -processors identical platform.

Lemma 12 *Let π denote any identical multiprocessors platform composed of m CPUs. Let S be any global, weakly work-conserving and FTP scheduler and let J be any set of n jobs sorted by decreasing S -priority, i.e., $J_1 >_S J_2 >_S \dots >_S J_n$. It holds $\forall k \in [1, m]$ and $\forall i \in [1, n]$ that*

$$\text{Work}_k^i = \begin{cases} \text{Work}_k^{i-1} + c_i & \text{if } k = \max \left\{ \underset{\ell \in [1, m]}{\text{argmin}} \{ \text{Work}_\ell^{i-1} \} \right\} \\ \text{Work}_k^{i-1} & \text{otherwise} \end{cases} \quad (8)$$

where $\text{Work}_k^0 = 0 \forall k$ by definition of the processed work.

Proof 10 *The proof directly follows from the definition of $\text{Work}_k^i \forall i, k$ and from the second condition of our definition of a weakly work-conserving scheduler (see Definition 8, page 6). Indeed, whenever a subset P of several CPUs idle (or complete a job) at the same time, S dispatches the waiting job (if any) with the highest priority to the CPU of P with the highest index (this is the reason for the condition “if k is the highest value of ℓ that minimizes Work_ℓ^{i-1} ”).*

Corollary 3 *An upper-bound $\overline{\text{idle}}_k$, $1 \leq k \leq m$, is given by the k^{th} element of the vector $\{\text{Work}_1^n, \text{Work}_2^n, \dots, \text{Work}_m^n\}$ sorted by non-decreasing order.*

Proof 11 *The proof directly follows from the definition of the processed work $\text{Work}_k^n, \forall k \in [1, m]$.*

Corollary 4 *The maximum makespan $\overline{\text{ms}}^{\text{ident}}(J, \pi, \mathcal{P})$ is given by $\overline{\text{idle}}_m$, where $\overline{\text{idle}}_m$ is determined as in Corollary 3.*

7.2 Accuracy of the upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi, \mathcal{P})$

In this section we prove that the upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi, \mathcal{P})$ is 1-competitive, i.e., *exact*—exact in the sense that it can actually be reached if every job executes for its WCET. Again, this is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET as we have to consider the worst-case scenario in which every old-mode task releases a job *exactly* upon the mode change request and all these jobs executes for their WCET during the transition.

For any transition from a given mode M^i to any other mode M^j , the knowledge of the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ and the fact that we proceed by simulation allow us to compute the *exact* maximum idle-instants simply by drawing the schedule of $\mathcal{J}_i^{\text{wc}}$ following \mathcal{P} and by measuring the idle-instants in this schedule. Using this approach, the measured upper-bound $\overline{\text{ms}}^{\text{ident}}(J, \pi, \mathcal{P})$ is nothing else but 1-competitive.

7.3 Validity tests for SM-MSO and AM-MSO

From Corollary 4, the *sufficient* validity test given by Test 1 (on page 9) can be rewritten as follows.

Validity Test 3 (SM-MSO, identical and FTP) *For any multi-mode real-time application τ and any identical platform π composed of m CPUs, the protocol SM-MSO is valid provided that, for every mode M^i ,*

$$\overline{\text{ms}}^{\text{ident}}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i) \leq \min_{j \neq i} \left\{ \min_{k=1}^{n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

where $\overline{\text{ms}}^{\text{ident}}(\mathcal{J}_i^{\text{wc}}, \pi)$ is defined as in Corollary 4, \mathcal{P}^i is obtained from the old-mode scheduler S^i and $\mathcal{J}_i^{\text{wc}}$ is defined as follows:

- ▷ $\mathcal{J}_i^{\text{wc}} \stackrel{\text{def}}{=} \{J_1, J_2, \dots, J_{n_i}\}$
- ▷ each job $J_k \in \mathcal{J}_i^{\text{wc}}$ has a processing time equal to the WCET C_k^i of task τ_k^i
- ▷ $\mathcal{J}_i^{\text{wc}}$ is sorted by decreasing S^i -priority.

Concerning the protocol AM-MSO, the upper-bounds $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{P}^i)$ (for all $1 \leq k \leq m$) determined in Corollary 3 can be used at line 10 of the validity algorithm given by Algorithm 10 on page 15).

8 Uniform platforms and FJP schedulers

8.1 Some useful observations

In this section, we show that the maximum makespan determination problem is highly counter-intuitive upon *uniform* platforms and the methods for solving this problem cannot be straightforwardly extended from those proposed for *identical* multiprocessor platforms. First, recall that the schedulers are assumed to be *strongly* work-conserving here since we focus on uniform platforms.

Observation 2 *For a given set of jobs, an intuitive idea for maximizing the makespan upon any m -processor uniform platform is to execute, at any time, the longest job upon the slowest CPU, i.e., the shorter the computation requirement of a job, the higher its priority. We name this priority assignment “Shortest Job First” (SJF). However, we can show by using the following example that this intuitive idea is erroneous, as SJF does not lead to the maximum makespan.*

Example 8 *Let us consider the set J of 4 jobs J_1, J_2, J_3, J_4 of respective processing times 4, 4, 16 and 22, and suppose that they are scheduled on the 2-processors uniform platform $\pi = [1, 2]$. The priority assignment SJF (i.e., $J_1 > J_2 > J_3 > J_4$) provides a makespan of 17.75 whereas the priority assignment $J_3 > J_1 > J_2 > J_4$ leads to a makespan of 19. Notice that the problem of determining in a polynomial time (i.e., without trying every priority assignment) a priority assignment leading to the maximum makespan remains an open question and is out of the scope of this study.*

Observation 3 *Another intuitive idea is to naively extend to uniform platforms the result (replicated below) of Corollary 2 on page 22, i.e., for any identical platform π composed of m CPUs, an upper-bound on the makespan is given by*

$$\overline{\text{ms}}^{\text{ident}}(J, \pi) \stackrel{\text{def}}{=} \begin{cases} c_n & \text{if } (n = m) \\ \frac{\sum_{i=1}^{n-1} c_i}{m} + c_n & \text{otherwise} \end{cases} \quad (9)$$

where c_i is assumed to be such that $c_i \geq c_{i-1} \forall i \in [2, n]$.

Upon *identical* platforms there is a sense in distinguishing the case $n = m$ from the case $n > m$, because the rem-jobs never migrate between CPUs during mode transitions. Therefore, in the particular case where $n = m$, the maximum makespan does *not* depend on the job priority assignment and can be determined *exactly* by $\overline{\text{ms}}^{\text{ident}}(J, \pi) = c_n$. In contrast, we can easily show that this property does not hold upon *uniform* platforms. That is, the maximum makespan in the case $n = m$ is *not* independent from the job priority assignment upon uniform platforms. This is shown through the following example.

Example 9 *Consider the uniform platform $\pi = [1, 2]$ and the two jobs J_1, J_2 of processing time 4 and 6, respectively. If $J_1 > J_2$ then J_1 completes on π_2 at time 2—time during which J_2 executes 2 execution units on π_1 —and J_2 completes on π_2 at time 4, thus leading to a makespan of 4. On the other hand, if $J_2 > J_1$ then J_2 completes on π_2 at time 3—time during which J_1 executes 3 execution units on π_1 —and J_1 completes on π_2 at time 3.5, thus leading to a makespan of 3.5. As a result, the maximum makespan in the case $n = m$ depends on the job priority assignment on uniform platforms and the case $n = m$ can no longer be distinguished from the case $m < n$.*

From the previous example, naively extending Expression 9 to uniform platforms yields the following “1-piece” expression¹⁰:

$$\overline{\text{ms}}_0^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{n-1} c_i}{s(1)} + \frac{c_n}{s_m} \quad (10)$$

Unfortunately, we show in the following example that this extension does not provide an upper-bound on the maximum makespan.

Example 10 *Let us consider the set J of 3 jobs J_1, J_2, J_3 of respective processing times 50, 80 and 99, and suppose that they are scheduled on the 3-processors uniform platform $\pi = [1, 2, 10]$. The maximum makespan is 20, reached using the job priority assignment $J_1 > J_2 > J_3$ (see Figure 18). On the other hand, Expression 10 yields $\overline{\text{ms}}_0^{\text{unif}}(J, \pi) = \frac{50+80}{13} + \frac{99}{10} = 19.9$. This approximation made by Expression 10 is illustrated in Figure 19. This simple example is much more important than what it seems to be at first blush and we will deeply examine its impacts in Section 8.4 (page 30).*

¹⁰recall that $s(1) \stackrel{\text{def}}{=} \sum_{i=1}^m s_i$

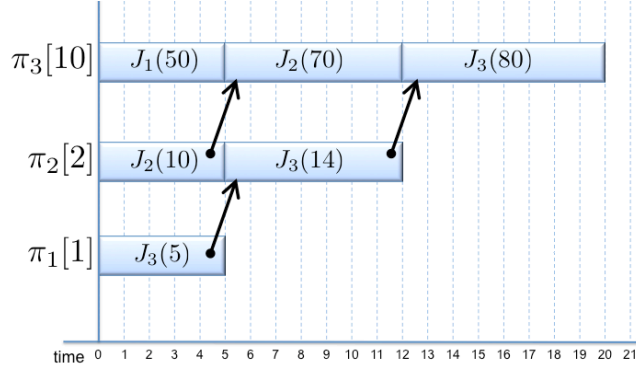


Figure 18: This picture depicts a priority assignment leading to a makespan of 20. The speed of each CPU is indicated into brackets next to its label. The numbers next to each job name J_i is the amount of work processed by J_i upon the allocated CPU. For instance, job J_1 executes 50 execution units from time 0 to 5 on CPU π_3 , leading to its label $J_1(50)$.

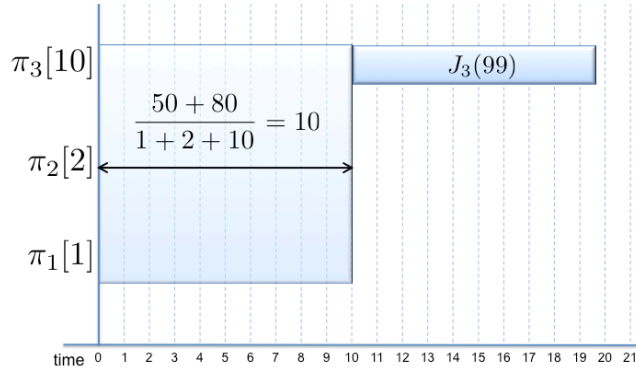


Figure 19: Approximation error made by Expression 10.

8.2 Upper-bounds $\overline{\text{idle}}_k(J, \pi)$ on the idle-instants

Once more but this time for any *uniform* platform π , we focus on determining a mathematical expression that provides an upper-bound $\overline{\text{idle}}_k(J, \pi)$ on the k^{th} idle-instant, $\forall k \in [1, m]$. For sake of clarity, the following two lemmas use the notations $\overline{\text{idle}}_k$ instead of $\overline{\text{idle}}_k(J, \pi)$ and similarly, the notation idle_k will be used to denote the *exact* value of the k^{th} idle-instant. First, Lemma 13 determines a *lower-bound* $\underline{\text{idle}}_k$ on each idle-instant idle_k , $1 \leq k \leq m$. Then, Lemma 14 determines an *upper-bound* $\overline{\text{idle}}_k$ on each idle-instant idle_k . Finally, Corollary 5 derives an upper-bound on the maximum makespan (recall that the maximum makespan is simply given by idle_m).

Lemma 13 (See [27]) *Let $\pi = [s_1, s_2, \dots, s_m]$ be any m -processors uniform platform such that $s_i \geq s_{i-1} \forall i$, $2 \leq i \leq m$. Let $J = \{J_1, J_2, \dots, J_n\}$ be any set of n jobs of respective processing times c_1, c_2, \dots, c_n such that $c_1 \leq c_2 \leq \dots \leq c_n$. Let S be the schedule of J upon π following any global, strongly work-conserving and FJP scheduler. A lower bound $\underline{\text{idle}}_k$ on each idle-instant idle_k ($1 \leq k \leq m$) in S is given by*

$$\underline{\text{idle}}_k \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{n-m+k} c_i}{s(1)} \quad (11)$$

Proof 12 *According to the definition of the idle-instants, at most $(m - k)$ jobs are not completed at time idle_k , meaning that at least $(n - m + k)$ jobs are already completed. Let J^{any} be any subset of J composed of r jobs, where $(n - m + k) \leq r \leq n$. Obviously, a lower bound t on the instant at which the r jobs of J^{any} are completed is given by*

$$t \stackrel{\text{def}}{=} \frac{\sum_{J_i \in J^{\text{any}}} c_i}{s(1)}$$

and since $c_1 \leq c_2 \leq \dots \leq c_n$, t is minimal if (i) the number of jobs in J^{any} is low as possible, i.e., $r = n - m + k$, and (ii) the processing time of each job of J^{any} is low as possible. As a result, t is minimum for $J^{\text{any}} = \{J_1, J_2, \dots, J_{n-m+k}\}$ and then yields a lower-bound for $\underline{\text{idle}}_k$.

Lemma 14 (See [27]) Using the same notations as in the previous lemma, an upper-bound $\overline{\text{idle}}_k$ on each idle-instant idle_k ($1 \leq k \leq m$) in S is given by

$$\overline{\text{idle}}_k \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n c_i - \sum_{i=1}^{k-1} \text{idle}_i \cdot s_i}{s(k)} \quad (12)$$

where $s(k) \stackrel{\text{def}}{=} \sum_{i=k}^m s_i$ (as defined in Expression 1, page 4).

Proof 13 From the “staircase” property derived from the definition of a strongly work-conserving scheduler on uniform platform (see page 5 for details) and from the fact that all the jobs are assumed to be synchronous at time 0, we know that CPU π_j becomes idle at time idle_j , $\forall j = 1, 2, \dots, m$. Let w_j ($1 \leq j \leq m$) denotes the amount of work executed on CPU π_j within $[0, \text{idle}_j]$, i.e., $w_j \stackrel{\text{def}}{=} \text{idle}_j \cdot s_j$. The proof is made by contradiction. Let ℓ be any integer in $[1, m]$ and suppose that $\text{idle}_\ell > \overline{\text{idle}}_\ell$. By definition of w_j , we know that

$$\sum_{j=1}^m w_j = \sum_{i=1}^n c_i \quad (13)$$

and from the definition of w_j we know that

$$\begin{aligned} \sum_{j=1}^m w_j &= \sum_{j=1}^m \text{idle}_j \cdot s_j \\ &= \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \sum_{j=\ell}^m (\text{idle}_j \cdot s_j) \end{aligned}$$

By definition of the idle-instants, it holds $\forall j \geq \ell$ that $\text{idle}_j \geq \text{idle}_\ell$. Therefore, replacing “ idle_j ” with “ idle_ℓ ” in the second term of the right-hand side of the above equality yields

$$\begin{aligned} \sum_{j=1}^m w_j &\geq \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \sum_{j=\ell}^m (\text{idle}_\ell \cdot s_j) \\ &\geq \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \text{idle}_\ell \cdot \sum_{j=\ell}^m s_j \end{aligned}$$

By hypothesis we have $\text{idle}_\ell > \overline{\text{idle}}_\ell$. Therefore, replacing idle_ℓ with $\overline{\text{idle}}_\ell$ in the right-hand side of the above inequality yields

$$\begin{aligned} \sum_{j=1}^m w_j &> \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \overline{\text{idle}}_\ell \cdot \sum_{j=\ell}^m s_j \\ &> \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \frac{\sum_{i=1}^n c_i - \sum_{i=1}^{\ell-1} \text{idle}_i \cdot s_i}{\sum_{i=\ell}^m s_i} \cdot \sum_{j=\ell}^m s_j \\ &> \sum_{j=1}^{\ell-1} (\text{idle}_j \cdot s_j) + \sum_{i=1}^n c_i - \sum_{i=1}^{\ell-1} \text{idle}_i \cdot s_i \\ &> \sum_{i=1}^n c_i + \sum_{j=1}^{\ell-1} ((\text{idle}_j - \underline{\text{idle}}_j) \cdot s_j) \end{aligned}$$

Since from Lemma 13 it holds that $\underline{\text{idle}}_i \leq \text{idle}_i \forall i = 1, 2, \dots, m$, it holds that

$$\sum_{j=1}^{\ell-1} ((\text{idle}_j - \underline{\text{idle}}_j) \cdot s_j) \geq 0$$

and thus

$$\sum_{j=1}^m w_j > \sum_{i=1}^n c_i$$

leading to a contradiction with Equality 13. The lemma follows.

Corollary 5 (See [27]) *Whatever the job priority assignment, an upper-bound $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ on the makespan is given by*

$$\overline{\text{ms}}_1^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{1}{s_m} \cdot \left(\sum_{i=1}^n c_i - \sum_{i=1}^{m-1} \text{idle}_i \cdot s_i \right) \quad (14)$$

Proof 14 *Since the makespan corresponds to the idle-instant idle_m , an upper-bound on the makespan is given by idle_m . Therefore, the proof is obtained by simply replacing k with m in Expression 12.*

8.3 Accuracy of the upper-bound $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$

In this section we prove that the upper-bound $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ is $\frac{s(1)}{s_m}$ -competitive, with the interpretation that the value returned by $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ is at most $\frac{s(1)}{s_m}$ times the exact value of the maximum makespan for any given set J of jobs and uniform platform π . Once again, this is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET as we have to consider the worst-case scenario in which every old-mode task releases a job exactly upon the mode change request and all these jobs executes for their WCET during the transition.

Lemma 15 *For any set J of jobs sorted by non-decreasing job processing time and any uniform platform $\pi = [s_1, s_2, \dots, s_m]$ with $s_i \geq s_{i-1} \forall i$, $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ is $\alpha_1(\pi)$ -competitive, where $\alpha_1(\pi) \stackrel{\text{def}}{=} \frac{s(1)}{s_m}$.*

Proof 15 *Recall from Expression 14 that*

$$\overline{\text{ms}}_1^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n c_i}{s_m} - \frac{\sum_{k=1}^{m-1} \left(\sum_{i=1}^{n-m+k} c_i \cdot s_k \right)}{s_m \cdot s(1)}$$

Let $\text{ms}(J, \pi)$ denote the exact makespan for any given set J of jobs and any uniform platform π . Since we do not have any mathematical expression for determining this exact makespan $\text{ms}(J, \pi)$, our analysis of $\alpha_1(\pi)$ is performed while considering a lower-bound $\widetilde{\text{ms}}(J, \pi)$ on the makespan rather than its exact value, i.e., $\alpha_1(\pi)$ is determined in such a manner that

$$\frac{\overline{\text{ms}}_1^{\text{unif}}(J, \pi)}{\widetilde{\text{ms}}(J, \pi)} \leq \alpha_1(\pi)$$

Obviously, we know that $\text{ms}(J, \pi) \geq \frac{\sum_{i=1}^n c_i}{s(1)}$ and this implies that $\widetilde{\text{ms}}(J, \pi) \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n c_i}{s(1)}$ is a lower-bound on the makespan. This yields

$$\frac{\overline{\text{ms}}_1^{\text{unif}}(J, \pi)}{\text{ms}(J, \pi)} \leq \frac{\overline{\text{ms}}_1^{\text{unif}}(J, \pi)}{\widetilde{\text{ms}}(J, \pi)}$$

and thus,

$$\begin{aligned} \frac{\overline{\text{ms}}_1^{\text{unif}}(J, \pi)}{\text{ms}(J, \pi)} &\leq \frac{\frac{\sum_{i=1}^n c_i}{s_m} - \frac{\sum_{k=1}^{m-1} \left(\sum_{i=1}^{n-m+k} c_i \cdot s_k \right)}{s_m \cdot s(1)}}{\frac{\sum_{i=1}^n c_i}{s(1)}} \\ &\leq \left(\frac{\sum_{i=1}^n c_i}{s_m} - \frac{\sum_{k=1}^{m-1} \left(\sum_{i=1}^{n-m+k} c_i \cdot s_k \right)}{s_m \cdot s(1)} \right) \cdot \frac{s(1)}{\sum_{i=1}^n c_i} \end{aligned} \quad (15)$$

$$\begin{aligned} &\leq \left(\frac{\sum_{i=1}^n c_i}{s_m} \right) \cdot \frac{s(1)}{\sum_{i=1}^n c_i} \\ &\leq \frac{s(1)}{s_m} \end{aligned} \quad (16)$$

Notice the important loss of accuracy that this inequality underwent when we ignored the term $\left(-\frac{\sum_{k=1}^{m-1} \left(\sum_{i=1}^{n-m+k} c_i \cdot s_k \right)}{s_m \cdot s(1)} \right)$ while passing from Inequality 15 to Inequality 16. The lemma follows.

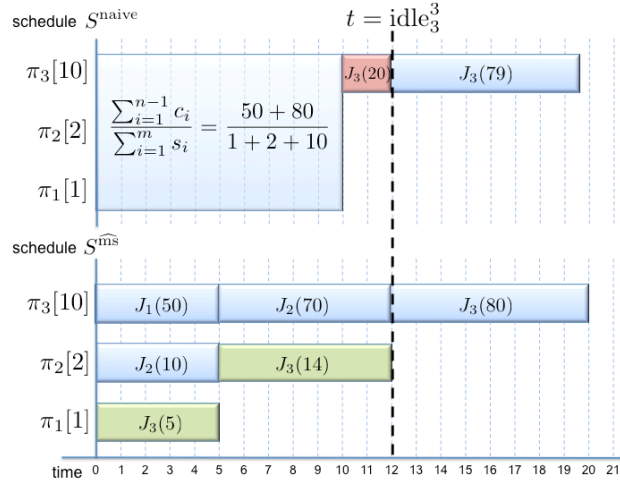


Figure 20: Example of schedule in which the makespan is larger than that returned by $\overline{ms}_0^{\text{unif}}(J, \pi)$.

8.4 Another analysis of the maximum makespan

In Example 10 on page 26, we have showed that the naive extension of $\overline{ms}^{\text{ident}}(J, \pi)$ (given by $\overline{ms}_0^{\text{unif}}(J, \pi)$ in Expression 10, page 26) does not provide an upper-bound on the maximum makespan considering uniform platforms. Essentially, in addition to refute the fact that $\overline{ms}_0^{\text{unif}}(J, \pi)$ provides an upper-bound on the maximum makespan, this example also refutes the *main concept behind the expression of $\overline{ms}^{\text{ident}}(J, \pi)$* . Indeed, in the expression of $\overline{ms}^{\text{ident}}(J, \pi)$, it can be easily shown that the term $\frac{\sum_{i=1}^{n-1} c_i}{m}$ is an upper-bound on the time at which J_n starts its execution, i.e., its *dispatching time*. Therefore, the whole expression can be interpreted as follows: upper-bound on the makespan = upper-bound on the dispatching time of $J_n + c_n$, where J_n is the (or any) job with the largest processing time. That is, this expression of $\overline{ms}^{\text{ident}}(J, \pi)$ is based on the intuition that the maximum makespan is reached when the longest job is dispatched as late as possible and executes for its WCET. This intuition has revealed to be true for the case of identical platforms, but not for the uniform case (as shown by Example 10)¹¹. The whole concept is not extendable to uniform platforms and in order to figure out the underlying cause, let us focus on Example 10.

Let S^{naive} and S^{ms} denote the two schedules depicted in Figure 20, issued from the approximation $\overline{ms}_0^{\text{unif}}(J, \pi)$ and from the priority assignment $J_1 > J_2 > J_3$ which leads to the maximum makespan, respectively. The reason why $\overline{ms}_0^{\text{unif}}(J, \pi)$ under-approximates the maximum makespan comes from the following fact: *if t denotes the instant at which job J_3 is dispatched to CPU π_3 in S^{ms} (here, $t = 12$), then during the time interval $[0, t]$, J_3 has executed a lower amount of execution units in the stairs of S^{ms} than upon π_3 in S^{naive}* . In other words the cumulated green areas in Figure 20 represent a lower amount of execution units than the red area. Indeed, J_3 executes $5 + 14 = 19$ execution units within $[0, t]$ in S^{ms} whereas it executes 20 execution units on π_3 in S^{naive} . As a result, the remaining processing time of J_3 at time t is higher in S^{ms} (here, 80) than in S^{naive} (here, 79), implying that J_3 completes later in S^{ms} than in S^{naive} . This is the reason why the expression $\overline{ms}_0^{\text{unif}}(J, \pi)$ does not provide the maximum makespan in the example above: on uniform platforms, the schedule in which any job J_i reaches its maximum completion time is not necessarily the schedule in which J_i is dispatched as late as possible.

Based on this fundamental observation, we propose and prove correct in [28] (pages 138–163 and 351–367) two additional upper-bounds $\overline{ms}_2^{\text{unif}}(J, \pi)$ and $\overline{ms}_3^{\text{unif}}(J, \pi)$ on the maximum makespan, considering uniform platforms and FJP schedulers. These upper-bounds are replicated below.

$$\overline{ms}_2^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{1}{s_m} \cdot \sum_{i=1}^n \left(c_i + s_1 \cdot \frac{\sum_{j=1}^{i-1} c_j}{s(1)} \right) \cdot K_{n-i} \quad (17)$$

where K_j is such that $\forall j$,

$$K_j \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } s_1 = s_m \text{ and } j = 0 \\ \left(1 - \frac{s_1}{s_m}\right)^j & \text{otherwise} \end{cases}$$

¹¹Indeed, we can also easily show that the term $\left(\sum_{i=1}^{n-1} c_i\right)/s(1)$ in Expression 10 is an upper-bound on the dispatching time of J_n and at that time J_n is dispatched to the fastest CPU π_m , leading to a WCET of $\frac{c_m}{s_m}$.

and

$$\overline{\text{ms}}_3^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{1}{s_m} \cdot \sum_{\ell=1}^n \left(c_\ell + \frac{s_x \cdot s_m \cdot \sum_{j=1}^{\ell-1} c_j}{s(1) \cdot \sum_{j=1}^x s_j} \right) \cdot H_{n-\ell} \quad (18)$$

where

$$x = \operatorname{argmin}_{i \in [1, m]} \left\{ \frac{s_i}{\sum_{j=1}^i s_j} \right\}$$

and H_j is such that $\forall j$,

$$H_j \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } s_x = \sum_{i=1}^x s_i \text{ and } j = 0 \\ \left(1 - \frac{s_x}{\sum_{i=1}^x s_i}\right)^j & \text{otherwise} \end{cases}$$

Each of these two upper-bounds is based on a distinct upper-bound on the amount of execution units that can be executed in the green areas (see Figure 20), and then derives an upper-bound on the completion time of every job, and finally on the makespan.

8.5 Validity tests for SM-MSO and AM-MSO

From Expressions 14, 17, 18, and Corollary 1, a *sufficient* validity test for the protocol SM-MSO can therefore be formalized as follows.

Validity Test 4 (SM-MSO, uniform and FJP) *For any multi-mode real-time application τ and any uniform platform $\pi = [s_1, s_2, \dots, s_m]$ composed of m CPUs, the protocol SM-MSO is valid provided that, for every mode M^i ,*

$$\overline{\text{ms}}_{\min}^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi) \leq \min_{j \neq i} \left\{ \min_{k=1}^{n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

where $\overline{\text{ms}}_{\min}^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi)$ is defined as $\overline{\text{ms}}_{\min}^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi) \stackrel{\text{def}}{=}$

$$\min \left\{ \overline{\text{ms}}_1^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi), \overline{\text{ms}}_2^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi), \overline{\text{ms}}_3^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi) \right\} \quad (19)$$

and $\overline{\text{ms}}_1^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi)$, $\overline{\text{ms}}_2^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi)$ and $\overline{\text{ms}}_3^{\text{unif}}(\mathcal{J}_i^{\text{wc}}, \pi)$ are defined as in Expressions 14, 17 and 18, respectively. This is performed considering the set $\mathcal{J}_i^{\text{wc}}$ composed of n_i jobs of processing time $C_1^i, C_2^i, \dots, C_{n_i}^i$ such that $C_j^i \geq C_{j-1}^i \forall j = 2, 3, \dots, n_i$.

Concerning the protocol AM-MSO, the upper-bounds $\overline{\text{idle}}_k(\mathcal{J}_i^{\text{wc}}, \pi)$ (for all $1 \leq k \leq m$) defined as in Lemma 14 can be used at line 10 of the validity algorithm given by Algorithm 10 (on page 15).

8.6 Simulation results

Because our analysis of the competitive factor did not lead to a constant α for the upper-bound $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ (as well as for the upper-bounds $\overline{\text{ms}}_2^{\text{unif}}(J, \pi)$ and $\overline{\text{ms}}_3^{\text{unif}}(J, \pi)$ as shown in [28]), this section reports on the results of simulations in order to quantify the precision of the three upper-bounds $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$, $\overline{\text{ms}}_2^{\text{unif}}(J, \pi)$ and $\overline{\text{ms}}_3^{\text{unif}}(J, \pi)$. These simulations are performed considering a single set J of jobs scheduled and multiple uniform platforms. We consider only a single set J of jobs for which the *exact* processing times are given in Table 5. We will explain below where these parameters are drawn from and why we consider only a single set of jobs rather than generating numerous job sets.

c_1	c_2	c_3	c_4	c_5
3896	3964	878	1378	2228
c_6	c_7	c_8	c_9	c_{10}
3612	1230	1232	1668	4672

Table 5: Processing times of the 10 jobs in J .

For experimental purposes, let us introduce the parameter λ_π defined in [18] for any m -processor uniform platform $\pi = [s_1, s_2, \dots, s_m]$,

$$\lambda_\pi \stackrel{\text{def}}{=} \max_{j=1}^m \left\{ \frac{\sum_{k=1}^{j-1} s_k}{s_j} \right\}$$

Informally speaking, this parameter λ_π measures the “degree” by which π differs from an identical multi-processor platform, i.e., its “degree of heterogeneity”. For any identical platform composed of m CPUs, it holds that $s_1 = s_2 = \dots = s_m$ and thus, $\lambda_\pi \stackrel{\text{def}}{=} \max_{j=1}^m \left\{ \frac{\sum_{k=1}^{j-1} s_k}{s_j} \right\}$ is maximum for $j = m$, leading to $\lambda_\pi \stackrel{\text{def}}{=} \frac{\sum_{k=1}^{m-1} s_k}{s_m} = m - 1$. The more homogeneous the platform π is, the closer to $(m - 1)$ is its corresponding λ_π . For instance, the uniform platform $\pi = [1, 500, 1000]$ has a corresponding $\lambda_\pi = \frac{501}{1000} \approx 0.5$ whereas $\lambda_\pi = \frac{501}{600} = 0.835$ for the uniform platform $\pi = [1, 500, 600]$ and $\lambda_\pi = \frac{1000}{600} \approx 1.67$ for the platform $\pi = [500, 500, 600]$. In short, $\lambda_\pi = (m - 1)$ if π is comprised of m identical CPUs and becomes progressively smaller as the speeds of the CPUs differ from each other by greater amounts.

The platform π considered in our simulations is composed of $m = 4$ CPUs for which we make their computing speed varying within $[1, 101]$ with an increment of 10. More precisely, we consider all possible combinations of the CPU speeds in the range $[1, 101]$ with an increment of 10, i.e., the first simulation is performed considering $\pi = [1, 1, 1, 1]$, the second simulation considers $\pi = [1, 1, 1, 11]$, the third one considers $\pi = [1, 1, 1, 21]$, and so on until reaching the speed assignment $\pi = [101, 101, 101, 101]$. For every speed assignment, we determine the corresponding parameter λ_π as well as the *exact* value $\text{ms}(J, \pi)$ of the maximum makespan. This exact maximum makespan $\text{ms}(J, \pi)$ is determined by building the schedule of J upon π for *every* job priority assignment and by retaining only the maximum generated makespan. This is a highly computational-intensive operation that requires the exhaustive enumeration of every possible job priority assignment. This is the reason why we consider only a single set J of jobs in our simulations. Indeed, according to this approach, our simulation process considers 11 different speeds for each CPU, leading to a total of $11^m = 11^4 = 14,641$ different platforms π . For each platform π , the computation of the exact makespan requires to generate the schedules derived from every job priority assignment. Since there are 10 jobs, the number of considered priority assignments is $10! = 3,628,800$. Multiplied by the number of platforms, this leads to 53,129,260,800 operations. Our simulations were performed on HYDRA, the Scientific Computer Configuration at the VUB/ULB Computing Centre, where we fully distributed the computations among 15 processors AMD Opteron dual-core @ 2.8GHz. Distributing the computations allowed us to complete the simulation in about 2 hours but unfortunately, the computation time grows exponentially with the number of CPUs and in a factorial manner with the number of jobs. For instance, considering 13 jobs would result in 91,169,811,532,800 operations, 14 jobs to approximately $20 \cdot 10^{15}$ operations, resulting in a computation time of about 82 years. The processing times of the jobs have been drawn from [15] where the authors present realistic parameters that concern the avionic domain. But since the number of operations of our algorithm is strongly restricted by the number of jobs, we arbitrarily selected 10 WCETs from these parameters.

For each speed assignment of the platform we computed the error $E_1^{\text{unif}}(J, \pi)$ corresponding to the difference (in percent) between $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ and $\text{ms}(J, \pi)$. Formally,

$$E_1^{\text{unif}}(J, \pi) \stackrel{\text{def}}{=} \frac{\overline{\text{ms}}_1^{\text{unif}}(J, \pi) - \text{ms}(J, \pi)}{\text{ms}(J, \pi)} \cdot 100$$

and in a similar way we also computed the errors $E_2^{\text{unif}}(J, \pi)$ and $E_3^{\text{unif}}(J, \pi)$.

The errors $E_1^{\text{unif}}(J, \pi)$, $E_2^{\text{unif}}(J, \pi)$ and $E_3^{\text{unif}}(J, \pi)$ are displayed in Figure 21 relative to the corresponding λ_π . The horizontal black line is the error “E.EXACT_MAKESPAN” of $\text{ms}(J, \pi)$ over the exact value of the maximum makespan. Obviously, this error is always 0. Also, for every speed assignment of π , we define the estimator $\overline{\text{ms}}_{\min}^{\text{unif}}(J, \pi)$ as in Expression 19 and its associated error $E_{\min}^{\text{unif}}(J, \pi)$. This error is displayed in Figure 22 relative to the corresponding λ_π . Finally, Table 6 provides the reader with some statistics issued from the simulation.

For obvious reason, the most accurate estimator (i.e., the most accurate upper-bound on the maximum makespan) is $\overline{\text{ms}}_{\min}^{\text{unif}}(J, \pi)$. As presented in Table 6, the most important error that we obtained for $\overline{\text{ms}}_{\min}^{\text{unif}}(J, \pi)$ is 22.89% and the minimal one is 1.57%. The average error is 10.44% with a squared distance of 5.78%. Hence, we believe that this is a promising path to go for more competitive bounds and for practical use. An open question remains however. For $\lambda_\pi \in [0, 2]$, we can see in Figure 21 that $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ is clearly lower than both $\overline{\text{ms}}_2^{\text{unif}}(J, \pi)$ and $\overline{\text{ms}}_3^{\text{unif}}(J, \pi)$, i.e., $\overline{\text{ms}}_{\min}^{\text{unif}}(J, \pi) = \overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ for $\lambda_\pi \in [0, 2]$. Within this interval $[0, 2]$, when the parameter λ_π reaches an integer value (here, 1 and 2), something happens that considerably improves the accuracy of $\overline{\text{ms}}_{\min}^{\text{unif}}(J, \pi)$. But up to now, we did not find any interpretation to that phenomenon.

9 Uniform platforms and FTP schedulers

This section follows the same reasoning as the one for identical platforms and FTP schedulers. For any transition from a specific mode M^i to any other mode M^j , the knowledge of the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ and the fact that the priorities are known beforehand enable us to compute the *exact* maximum idle-instants idle_k , $1 \leq k \leq m$, simply by simulating the scheduling of the critical rem-job set and by measuring the idle-instants idle_k , $1 \leq k \leq m$, in that schedule (from Corollary 1 presented on page 19). Thus, each idle-instant idle_k measured in the

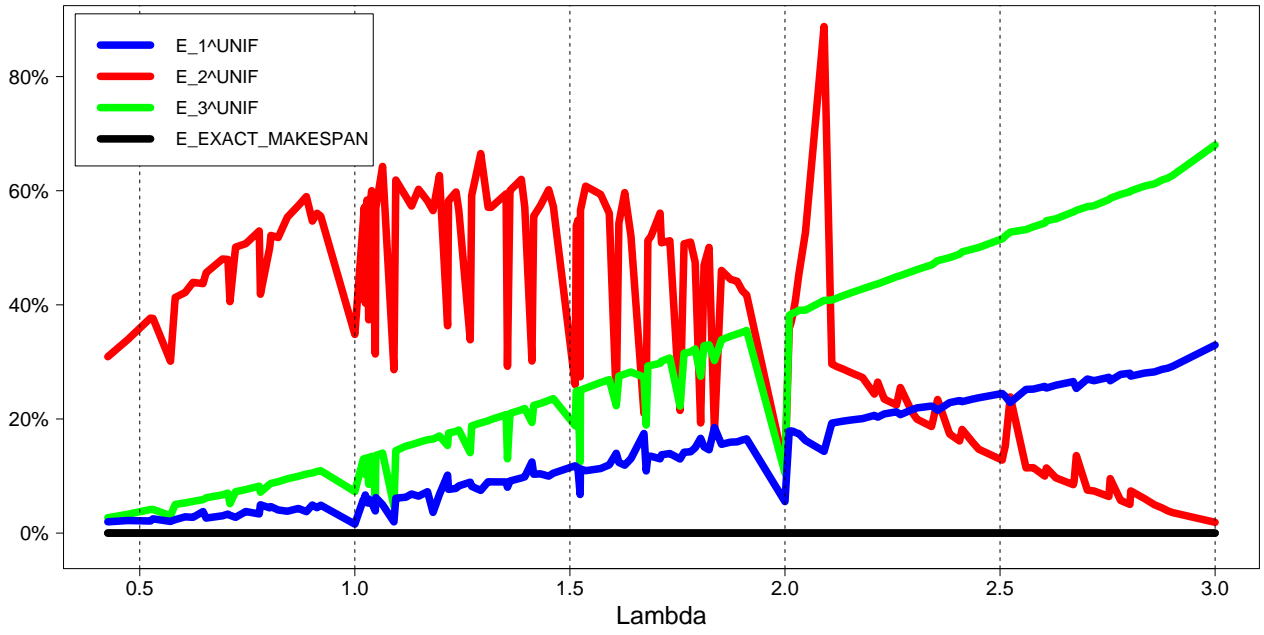


Figure 21: The three estimation errors $E_1^{\text{unif}}(J, \pi)$, $E_2^{\text{unif}}(J, \pi)$ and $E_3^{\text{unif}}(J, \pi)$ displayed relative to the corresponding λ_π .

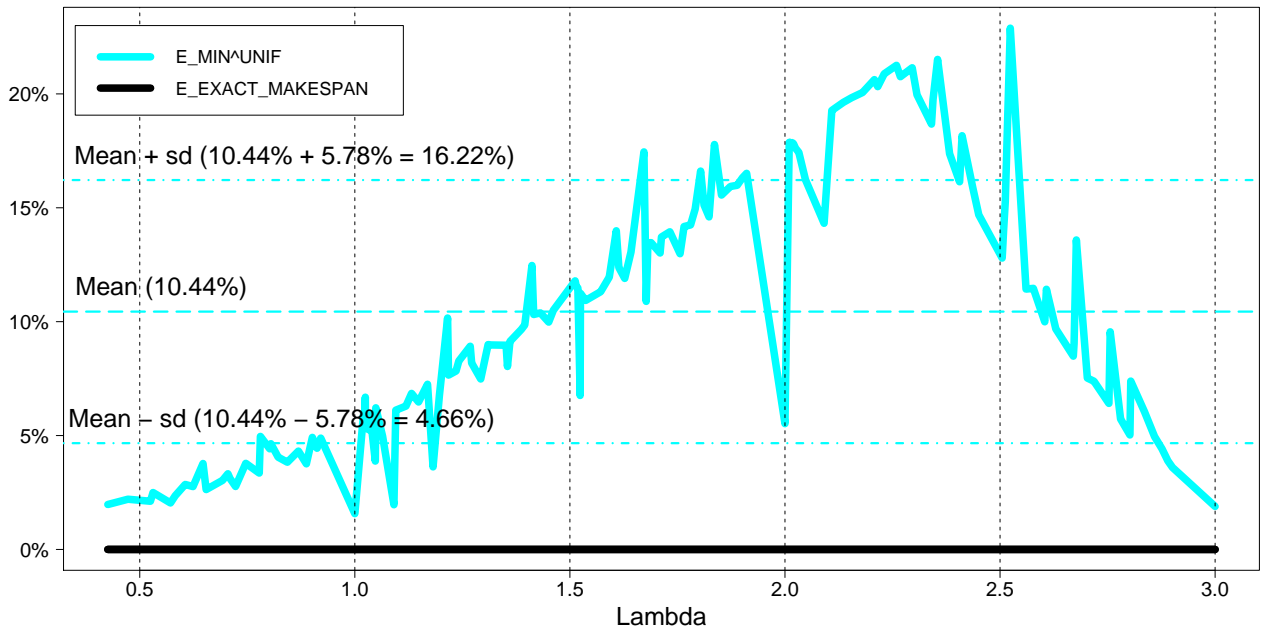


Figure 22: The estimation error $E_{\text{min}}^{\text{unif}}(J, \pi)$ displayed relative to the corresponding λ_π .

	$E_1^{\text{unif}}(J, \pi)$	$E_2^{\text{unif}}(J, \pi)$	$E_3^{\text{unif}}(J, \pi)$	$E_{\min}^{\text{unif}}(J, \pi)$
Min.	1.57%	1.89%	2.7%	1.57%
1st Qu.	6%	21.74%	13.28%	5.3%
Median	12.72%	41.07%	27.11%	9.92%
Mean	13.68%	37.91%	29.25%	10.44%
3rd Qu.	20.72%	55.5%	43.99%	15.08%
Max.	32.96%	88.78%	68.01%	22.89%
Variance	69.76	359.37	320.47	33.36
SD ¹²	8.35%	18.96%	17.9%	5.78%
Bias	42.35%	133.69%	94.05%	26.93%
MSE ¹³	1863.25	18233.15	9165.97	758.43

Table 6: Statistics issued from the simulation

schedule of the critical rem-job set is an upper-bound on the idle-instants idle_k in the schedule derived from any other set of rem-jobs. In conclusion, FTP schedulers enable us to determine the *exact*¹⁴ maximum idle-instants idle_k , $1 \leq k \leq m$, rather than over-approximating them (as done for the FJP schedulers).

9.1 Upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ on the idle-instants

Lemma 16 provides the exact values of $\text{idle}_j(J_i, \pi, \mathcal{P}) \forall j \in [1, m], i \in [1, n]$ and $\forall \mathcal{P}$, assuming that every job J_i executes for its WCET. However, in this particular case of FTP scheduler, we redefine the idle-instants $\text{idle}_j(J_i, \pi, \mathcal{P})$ as follows.

Definition 17 (Idle-instant $\text{idle}_j(J_i, \pi, \mathcal{P})$) *If S^i denotes the schedule upon π of only the jobs with a higher (or equal) priority than J_i according to \mathcal{P} , then $\text{idle}_j(J_i, \pi, \mathcal{P})$ is the earliest instant in S^i at which at least j CPUs idle.*

The only difference w.r.t. the previous one resides in the “higher (or equal) priority than (...)”. The reason for this redefinition is that, with the previous one, it was not possible to express the idle-instants $\text{idle}_j(J, \pi, \mathcal{P})$ (for $j = 1, \dots, m$) as in Definition 13 (page 14). Indeed, these idle-instants $\text{idle}_j(J, \pi, \mathcal{P})$ consider that every job of J are scheduled while the previous definition of the idle-instants $\text{idle}_j(J_i, \pi, \mathcal{P})$ requires a job index i and considers that only the jobs with a higher priority than J_i are scheduled. Thereby, this previous definition always excludes the job J_i in the computation of the idle-instants. Now, thanks to this new definition, the idle-instants $\text{idle}_j(J, \pi, \mathcal{P})$ (for $j = 1, \dots, m$) can be expressed by $\text{idle}_j(J_{\text{low}}, \pi, \mathcal{P})$, where J_{low} is the lowest priority job according to \mathcal{P} . Once again, we use in Corollary 6 the notations idle_j^i to refer to the idle-instants $\text{idle}_j(J_i, \pi, \mathcal{P})$ defined as in Definition 17.

Lemma 16 (See [27]) *Let $\pi = [s_1, s_2, \dots, s_m]$ denote any uniform multiprocessor platform composed of m CPUs and assume that $s_i \geq s_{i-1}, \forall i = 2, 3, \dots, m$. Let $J = \{J_1, J_2, \dots, J_n\}$ be any set of n jobs, all released at time $t = 0$, with respective computation time c_1, c_2, \dots, c_n . Let \mathcal{S} denote any global, FTP and strongly work-conserving scheduler and suppose that J is sorted by decreasing \mathcal{S} -priority, i.e., $J_i >_{\mathcal{S}} J_{i+1}$. If these jobs are scheduled by \mathcal{S} upon π , then idle_j^i is inductively defined as follows:*

Initialization:

$$\begin{aligned} \forall 1 \leq j \leq m : \quad \text{idle}_j^0 &\stackrel{\text{def}}{=} 0 \\ \forall 1 \leq i \leq n : \quad \text{idle}_{m+1}^i &\stackrel{\text{def}}{=} \infty \end{aligned}$$

Iteration:

for ($i = 1$ to n)
for ($j = m$ to 1)

$$\text{idle}_j^i \stackrel{\text{def}}{=} \begin{cases} \text{idle}_j^{i-1} & \text{if } \text{idle}_j^{i-1} = \text{idle}_{j+1}^{i-1} \\ \text{idle}_{j+1}^{i-1} & \text{else if } c_i \geq \sum_{k=1}^j (\text{idle}_{k+1}^{i-1} - \text{idle}_k^{i-1}) \cdot s_k \\ \text{idle}_j^{i-1} + \frac{(c_i - \sum_{k=1}^{j-1} (\text{idle}_{k+1}^{i-1} - \text{idle}_k^{i-1}) \cdot s_k)}{s_j} & \text{otherwise} \end{cases} \quad (20)$$

¹⁴Exact in the sense that this value is actually reached if every job executes for its WCET.

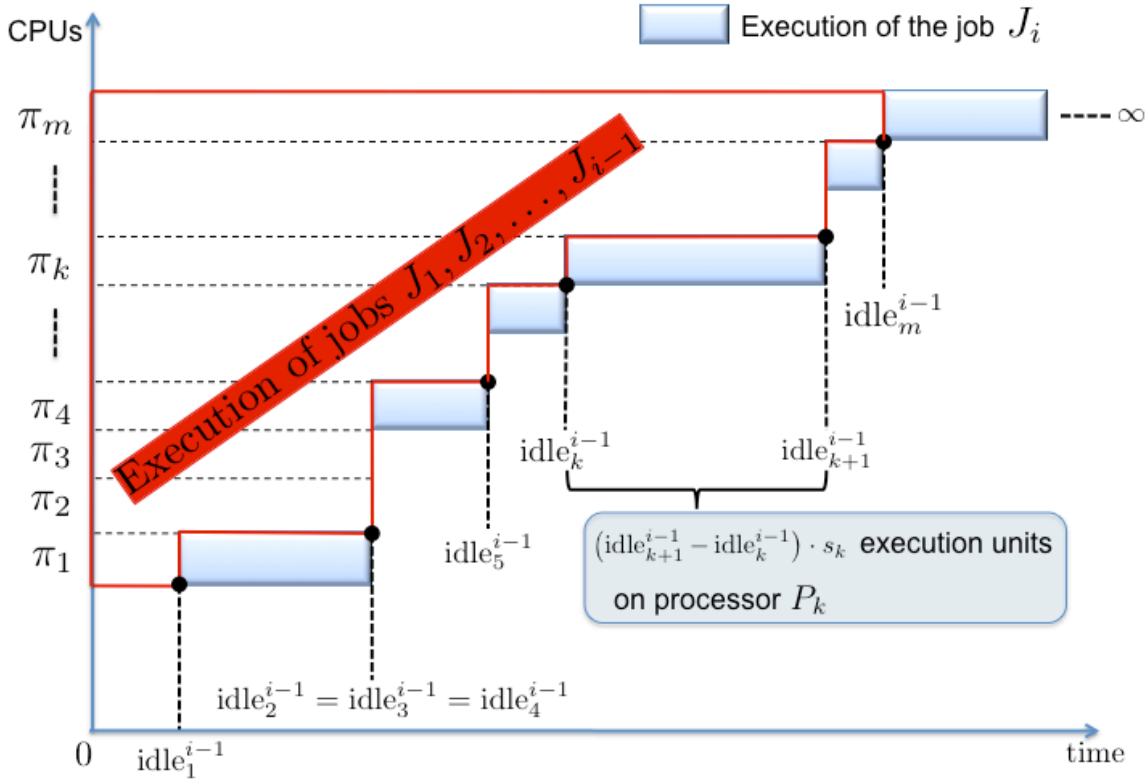


Figure 23: Staircase defined by the idle_j^{i-1}

Proof 16 Initially, the m CPUs are idle and thus, $\text{idle}_j^0 = 0, \forall j, 1 \leq j \leq m$. We find convenient to define $\text{idle}_{m+1}^i \stackrel{\text{def}}{=} \infty, \forall i$, which means that we have at most m CPUs available. In the following, we prove the correctness of the value of $\text{idle}_j^i (\forall j, 1 \leq j \leq m)$ assuming that idle_j^{i-1} are defined ($\forall j \leq m+1$). The idle-instants idle_j^{i-1} define a staircase as illustrated in Figure 23 for the scheduling of jobs J_1, \dots, J_{i-1} . Thus, job J_i can only progress into the blue areas and two cases have to be distinguished:

Case 1 $\text{idle}_j^{i-1} = \text{idle}_{j+1}^{i-1}$, meaning that at least one CPU faster than π_j becomes available at time idle_j^{i-1} (the blue area on CPU π_j is void in that case). This situation is depicted in Figure 23 where $\text{idle}_2^{i-1} = \text{idle}_3^{i-1} = \text{idle}_4^{i-1}$ and the blue area is void on CPUs π_2 and π_3 . In this kind of situation, the job J_i is executed (if not completed) upon a faster CPU and the first instant at which at least j CPUs idle remains unchanged after having scheduled the job J_i , i.e., $\text{idle}_j^i = \text{idle}_j^{i-1}$.

Case 2 Otherwise, J_i is dispatched to CPU π_j at instant idle_j^{i-1} and keeps executing on π_j as long as (i) no faster CPUs become idle or (ii) J_i completes. In the first case, J_i executes on π_j until the next idle-instant idle_{j+1}^{i-1} , leading to the first sub-case $\text{idle}_j^i = \text{idle}_{j+1}^{i-1}$. In the second case, J_i executes on CPU π_j but completes before time idle_{j+1}^{i-1} . Thus, the idle-instant idle_j^i is the instant idle_j^{i-1} at which J_i was dispatched to π_j plus its remaining processing time on CPU π_j at time idle_j^{i-1} . Since $\sum_{k=1}^j (\text{idle}_k^{i-1} - \text{idle}_{k-1}^{i-1}) \cdot s_k$ corresponds to the amount of work that J_i has executed in the interval of time $[0, \text{idle}_j^{i-1}]$, its remaining processing time on CPU π_j at time idle_j^{i-1} is given by $\frac{c_i - \sum_{k=1}^j (\text{idle}_k^{i-1} - \text{idle}_{k-1}^{i-1}) \cdot s_k}{s_j}$, leading to the second sub-case.

Corollary 6 (See [27]) The maximum idle-instant $\overline{\text{idle}}_k(J, \pi, \mathcal{P}) (\forall k \in [1, m])$ is given by idle_k^n computed as in Lemma 16.

Corollary 7 The maximum makespan $\overline{\text{ms}}_1^{\text{unif}}(J, \pi)$ is given by idle_m^n computed as in Lemma 16.

9.2 Validity tests for SM-MSO and AM-MSO

From Corollary 7, a sufficient validity test for the protocol SM-MSO can therefore be formalized as follows.

Validity Test 5 (SM-MSO, uniform and FTP) For any multi-mode real-time application τ and any identical platform π composed of m CPUs, the protocol SM-MSO is valid provided that, for every mode M^i ,

$$\text{idle}_m^{n_i}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{S}^i) \leq \min_{j \neq i} \left\{ \min_{k=1}^{n_j} \left\{ \mathcal{D}_k^j(M^i) \right\} \right\}$$

where $\text{idle}_m^{n_i}(\mathcal{J}_i^{\text{wc}}, \pi, \mathcal{S}^i)$ is computed as $\text{idle}_m^{n_i}$ in Lemma 16, considering the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ composed of n_i jobs J_1, J_2, \dots, J_{n_i} of respective processing time $C_1^i, C_2^i, \dots, C_{n_i}^i$ and such that $\mathcal{J}_i^{\text{wc}}$ is sorted by decreasing \mathcal{S}^i -priority.

Similarly, the upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ (where $1 \leq k \leq m$ and \mathcal{P} corresponds to the job priority assignment of the old-mode scheduler \mathcal{S}^i) determined in Lemma 16 can be used at line 10 of the validity algorithm of AM-MSO (see Algorithm 10 page 15), as long as these upper-bounds are computed while assuming the critical rem-job set $\mathcal{J}_i^{\text{wc}}$ for the transitions from every mode M^i .

10 Conclusion and open problems

In this paper, we addressed the scheduling problem of multi-mode real-time applications upon identical and uniform multiprocessor platforms. We assumed that every mode of the application was scheduled by following a global and Fixed-Task-Priority or Fixed-Job-Priority scheduler. Under these assumptions, we proposed two protocols for managing every transition between every pair of modes of the system, namely SM-MSO and AM-MSO. For both protocols, we established validity tests that allow the system designer to predict whether the given application can meet all the expected timing requirements upon the given platform. We prove the correctness of our schedulability analyses by extending the theory about the makespan determination problem.

In our future work, we aim at taking into account mode-independent tasks, i.e., tasks whose the periodic (or sporadic) activation pattern is not affected by the mode changes. Moreover, instead of scheduling the rem-jobs by using the scheduler of the old-mode during the transitions, it could be better, in term of the enablement delays applied to the new-mode tasks, to propose a *dedicated priority assignment* which meets the deadline of every rem-job, while minimizing the makespan. To the best of our knowledge, the problem of minimizing the makespan while meeting job deadlines is not yet addressed in the literature and remains open. Table 7 outlines a brief overview of all different problems, considering the task and platform model introduced in this paper. For each problem, we indicated either the reference(s) where solutions have been proposed or T.W. (This Work) or F.W. (Future Work) or O.P. (Open Problem).

References

- [1] ANDERSSON, B. *Static-priority scheduling on multiprocessors*. PhD thesis, Chalmers University of Technology, 2003.
- [2] ANDERSSON, B. Uniprocessor EDF scheduling with mode change. In *Proceedings of the 12th International Conference on Principles of Distributed Systems* (Berlin, Heidelberg, 2008), OPODIS'08, Springer-Verlag, pp. 572–577.
- [3] BAILEY, C. M. Hard real-time operating system kernel. investigation of mode change. Tech. rep., Task 14 Deliverable on ESTSEC Contract 9198/90/NL/SF, British Aerospace Systems Ltd., 1993.
- [4] BAKER, T. P. Multiprocessor EDF and deadline monotonic schedulability analysis. In *Proceedings of the 24th IEEE International Real-Time Systems Symposium (RTSS '03)* (Washington, DC, USA, 2003), IEEE Computer Society, pp. 120–129.
- [5] BAKER, T. P., AND BARUAH, S. K. An analysis of global EDF schedulability for arbitrary-deadline sporadic task systems. *Real-Time Systems* 43, 1 (2009), 3–24.
- [6] BAKER, T. P., AND CIRINEI, M. Brute-force determination of multiprocessor schedulability for sets of sporadic hard-deadline tasks. In *Proceedings of the 11th international conference on Principles of distributed systems* (Berlin, Heidelberg, 2007), OPODIS'07, Springer-Verlag, pp. 62–75.
- [7] BARUAH, S. An improved global EDF schedulability test for uniform multiprocessors. In *Proceedings of the 16th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS '10)* (Los Alamitos, CA, USA, 2010), IEEE Computer Society, pp. 184–192.

Protocols without periodicity			
Protocol	Platform	Scheduler	Existing results
Synchronous	identical	FJP	[30, 31, 28], T.W.
Synchronous	identical	FTP	[30, 31, 28], T.W.
Synchronous	uniform	FJP	[27, 28], T.W.
Synchronous	uniform	FTP	[27, 28], T.W.
Asynchronous	identical	FJP	[31, 28], T.W.
Asynchronous	identical	FTP	[31, 28], T.W.
Asynchronous	uniform	FJP	[27, 28], T.W.
Asynchronous	uniform	FTP	[27, 28], T.W.
Protocols with periodicity			
Protocol	Platform	Scheduler	Existing results
Synchronous	identical	FJP	[29], F.W.
Synchronous	identical	FTP	[29], F.W.
Synchronous	uniform	FJP	F.W.
Synchronous	uniform	FTP	F.W.
Asynchronous	identical	FJP	O.P.
Asynchronous	identical	FTP	O.P.
Asynchronous	uniform	FJP	O.P.
Asynchronous	uniform	FTP	O.P.

Table 7: State-of-the-art at a glance.

- [8] BARUAH, S., AND ANDERSON, J. Energy-aware implementation of hard-real-time systems upon multiprocessor platform. In *Proceedings of the 16th International Conference on Parallel and Distributed Computing Systems* (August 2003), pp. 430–435.
- [9] BARUAH, S., AND ANDERSON, J. Energy-efficient synthesis of periodic task systems upon identical multiprocessor platforms. In *Proceedings of the Twenty-Fourth International Conference on Distributed Computing Systems* (Tokyo, Japan, March 2004), IEEE Computer Society Press, pp. 428–435.
- [10] BARUAH, S., AND BAKER, T. Global EDF schedulability analysis of arbitrary sporadic task systems. In *Proceedings of the 2008 Euromicro Conference on Real-Time Systems (ECRTS '08)* (Washington, DC, USA, 2008), IEEE Computer Society, pp. 3–12.
- [11] BARUAH, S., AND FISHER, N. Global deadline-monotonic scheduling of arbitrary-deadline sporadic task systems. In *Proceedings of the 11th international conference on Principles of distributed systems (OPODIS '07)* (Guadeloupe, French West Indies, 2007), Springer-Verlag, pp. 204–216.
- [12] BARUAH, S., AND GOOSSENS, J. Deadline monotonic scheduling on uniform multiprocessors. In *Proceedings of the 12th International Conference on Principles of Distributed Systems (OPODIS '08)* (Berlin, Heidelberg, 2008), Springer-Verlag, pp. 89–104.
- [13] BARUAH, S., AND GOOSSENS, J. The EDF scheduling of sporadic task systems on uniform multiprocessors. In *Proceedings of the 2008 Real-Time Systems Symposium (RTSS '08)* (Washington, DC, USA, 2008), IEEE Computer Society, pp. 367–374.
- [14] BARUAH, S. K., AND GOOSSENS, J. Rate-monotonic scheduling on uniform multiprocessors. *IEEE Transactions on Computers* 52, 7 (2003), 966–970.
- [15] BATE, I. J. *Scheduling and Timing Analysis for Safety Critical Real-Time Systems*. PhD thesis, University of York, November 1998.
- [16] BERTOONA, M., CIRINEI, M., AND LIPARI, G. Improved schedulability analysis of EDF on multiprocessor platforms. In *Proceedings of the 17th Euromicro Conference on Real-Time Systems (ECRTS '05)* (Washington, DC, USA, 2005), IEEE Computer Society, pp. 209–218.

- [17] CUCU-GROSJEAN, L., AND GOOSSENS, J. Predictability of fixed-job priority schedulers on heterogeneous multiprocessor real-time systems. *Information Processing Letters* 110, 10 (2010), 399–402.
- [18] FUNK, S., GOOSSENS, J., AND BARUAH, S. On-line scheduling on uniform multiprocessors. In *Proceedings of the 22nd IEEE Real-Time Systems Symposium (RTSS '01)* (Washington, DC, USA, 2001), IEEE Computer Society, pp. 183–192.
- [19] GAREY, M. R., AND JOHNSON, D. S. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1990.
- [20] GOOSSENS, J., FUNK, S., AND BARUAH, S. Real-time scheduling on uniform multiprocessors. In *Proceedings of the 10th international conference on real-time systems* (Paris France, March 2002), In Teknea, editors, pp. 189–204.
- [21] GOYAL, S. V. 15-854: Approximations Algorithms, Lecturer: R. Ravi, Topic: Greedy Algorithms: Minimizing Makespan, Multiway Cut, September 2005.
- [22] HA, R. *Validating Timing Constraints in Multiprocessor and Distributed Systems*. PhD thesis, Department of Computer Science, University of Illinois at Urbana-Champaign, 1995.
- [23] HA, R., AND LIU, J. W. Validating timing constraints in multiprocessor and distributed real-time systems. Tech. rep., Department of Computer Science, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 1993.
- [24] HA, R., AND LIU, J. W. S. Validating timing constraints in multiprocessor and distributed real-time systems. In *Proceedings of the 14th IEEE International Conference on Distributed Computing Systems* (Los Alamitos, CA, USA, 1994), IEEE Computer Society Press, pp. 162–171.
- [25] HENIA, R., AND ERNST, R. Scenario aware analysis for complex event models and distributed systems. In *Proceedings of the 28th IEEE International Real-Time Systems Symposium (RTSS '07)* (Washington, DC, USA, 2007), IEEE Computer Society, pp. 171–180.
- [26] LIU, C. L., AND LAYLAND, J. W. Scheduling algorithms for multiprogramming in a hard-real-time environment. *Journal of ACM* 20, 1 (1973), 46–61.
- [27] MEUMEU YOMSI, P., NELIS, V., AND GOOSSENS, J. Scheduling multi-mode real-time systems upon uniform multiprocessor platforms. In *The 15th IEEE International Conference on Emerging Technologies and Factory Automation* (2010), IEEE Computer Society Press.
- [28] NELIS, V. *Energy-Aware Real-Time Scheduling in Multiprocessor Embedded Systems*. PhD thesis, Université Libre de Bruxelles, 2010.
- [29] NELIS, V., ANDERSSON, B., AND GOOSSENS, J. A synchronous transition protocol with periodicity for global scheduling of multimode real-time systems on multiprocessors. In *The 30th IEEE Real-Time Systems Symposium (RTSS) Work-in-progress session* (Washington D.C. USA, December 2009), D. Zhu, Ed., pp. 13–16.
- [30] NELIS, V., AND GOOSSENS, J. Mode change protocol for multi-mode real-time systems upon identical multiprocessors. In *Proceedings of the 29th IEEE Real-Time Systems Symposium (Work in Progress session - RTSS08-WiP)* (Barcelona Spain, December 2008), pp. 9–12.
- [31] NELIS, V., GOOSSENS, J., AND ANDERSSON, B. Two protocols for scheduling multi-mode real-time systems upon identical multiprocessor platforms. In *21st Euromicro Conference on Real-Time Systems (ECRTS'09)* (Dublin Ireland, July 2009), IEEE Computer Society, pp. 151–160.
- [32] PEDRO, P. *Schedulability of mode changes in flexible real-time distributed systems*. PhD thesis, University of York, Department of Computer Science, 1999.
- [33] PEDRO, P., AND BURNS, A. Schedulability analysis for mode changes in flexible real-time systems. In *Proceedings of the 10th Euromicro Workshop on Real-Time Systems* (Los Alamitos, CA, USA, 1998), IEEE Computer Society, pp. 172–179.
- [34] REAL, J., AND CRESPO, A. Mode change protocols for real-time systems: A survey and a new proposal. *Real-Time Systems* 26, 2 (2004), 161–197.
- [35] SHA, L., RAJKUMAR, R., LEHOCZKY, J., AND RAMAMRITHAM, K. Mode change protocols for priority-driven preemptive scheduling. *Real-Time Systems* 1 (1989), 243–264.

- [36] SHA, L., SHA, L., RAJKUMAR, R., RAJKUMAR, R., LEHOCZKY, J., LEHOCZKY, J., RAMAMRITHAM, K., AND RAMAMRITHAM, K. Mode change protocols for priority-driven preemptive scheduling. *Real-Time Systems 1* (1988), 243–264.
- [37] STOIMENOV, N., PERATHONER, S., AND THIELE, L. Reliable mode changes in real-time systems with fixed priority or EDF scheduling. In *Proceedings of the Conference on Design, Automation and Test in Europe (DATE '09)* (2009), pp. 99–104.
- [38] TINDELL, K., AND ALONSO, A. A very simple protocol for mode changes in priority preemptive systems. Tech. rep., Universidad Politécnica de Madrid, 1996.
- [39] TINDELL, K., BURNS, A., AND WELLINGS, A. J. Mode changes in priority pre-emptively scheduled systems. In *Proceedings of the 13th Real-Time Systems Symposium* (Phoenix, Arizona, 1992), pp. 100–109.